

The clustering of DESI-like luminous red galaxies using photometric redshifts

Rongpu Zhou¹,^{1,2,3}★ Jeffrey A. Newman,^{1,2} Yao-Yuan Mao^{1,2,4}†, Aaron Meisner,⁵ John Moustakas,⁶ Adam D. Myers,⁷ Abhishek Prakash,⁸ Andrew R. Zentner^{1,2}, David Brooks,⁹ Yutong Duan^{1,3,10}, Martin Landriau,³ Michael E. Levi,³ Francisco Prada¹¹ and Gregory Tarle¹²

¹Department of Physics and Astronomy, University of Pittsburgh, 3941 O'Hara Street, Pittsburgh, PA 15260, USA

²Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC), USA

³Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

⁴Department of Physics and Astronomy, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA

⁵NSF's National Optical-Infrared Astronomy Research Laboratory, 950 N. Cherry Avenue, Tucson, AZ 85719, USA

⁶Department of Physics and Astronomy, Siena College, 515 Loudon Road, Loudonville, NY 12211, USA

⁷Department of Physics & Astronomy, University of Wyoming, 1000 E. University, Dept. 3905, Laramie, WY 82071, USA

⁸Infrared Processing and Analysis Center, California Institute of Technology, MC 100-22, Pasadena, CA 91125, USA

⁹Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT, UK

¹⁰Physics Department, Boston University, 590 Commonwealth Avenue, Boston, MA 02215, USA

¹¹Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, s/n, E-18008 Granada, Spain

¹²Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA

Accepted 2020 November 30. Received 2020 November 30; in original form 2020 January 17

ABSTRACT

We present measurements of the redshift-dependent clustering of a DESI-like luminous red galaxy (LRG) sample selected from the Legacy Survey imaging data set, and use the halo occupation distribution (HOD) framework to fit the clustering signal. The photometric LRG sample in this study contains 2.7 million objects over the redshift range of $0.4 < z < 0.9$ over 5655 deg^2 . We have developed new photometric redshift (photo- z) estimates using the Legacy Survey DECam and *WISE* photometry, with $\sigma_{\text{NMAD}} = 0.02$ precision for LRGs. We compute the projected correlation function using new methods that maximize signal-to-noise ratio while incorporating redshift uncertainties. We present a novel algorithm for dividing irregular survey geometries into equal-area patches for jackknife resampling. For a five-parameter HOD model fit using the MultiDark halo catalogue, we find that there is little evolution in HOD parameters except at the highest redshifts. The inferred large-scale structure bias is largely consistent with constant clustering amplitude over time. In an appendix, we explore limitations of Markov chain Monte Carlo fitting using stochastic likelihood estimates resulting from applying HOD methods to N -body catalogues, and present a new technique for finding best-fitting parameters in this situation. Accompanying this paper, we have released the Photometric Redshifts for the Legacy Surveys catalogue of photo- z 's obtained by applying the methods used in this work to the full Legacy Survey Data Release 8 data set. This catalogue provides accurate photometric redshifts for objects with $z < 21$ over more than $16\,000 \text{ deg}^2$ of sky.

Key words: galaxies: distances and redshifts – galaxies: evolution – galaxies: haloes – large-scale structure of Universe.

1 INTRODUCTION

Galaxy surveys over the past two decades, such as the Sloan Digital Sky Survey (SDSS; Gunn et al. 1998; York 2000), Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013), and DEEP2 (Newman et al. 2013), have enabled remarkable advancement in our understanding of the Universe, and more recent and ongoing surveys such as the Dark Energy Survey (DES; The Dark Energy Survey Collaboration 2005) and the Hyper Suprime-Cam survey (Aihara et al. 2018b) are providing more stringent constraints on models of cosmology and galaxy evolution.

The Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration 2016a, b) is a next-generation galaxy redshift survey that will produce an unprecedented three-dimensional map of the Universe and shed light on the expansion history of the Universe and the nature of dark energy. An important class of DESI spectroscopic targets are the luminous red galaxies (LRGs). The high large-scale structure bias of the LRGs make them ideal tracers for the underlying matter distribution, and they have been used to efficiently measure the baryon acoustic oscillation (BAO) signal (e.g. Eisenstein et al. 2005; Alam et al. 2017). The DESI survey, with much larger survey volume and higher number density, will enable more accurate BAO measurements.

In this paper, we present small-scale ($\lesssim 20 h^{-1} \text{ Mpc}$) galaxy clustering analysis of a set of DESI-like LRGs. Although the small-scale clustering of LRGs from other programs have been studied

* E-mail: rongpuzhou@lbl.gov

† NHFP Einstein Fellow.

previously (e.g. Zheng et al. 2008; White et al. 2011; Zhai et al. 2017), these samples are either at lower redshifts or are much sparser than the DESI LRG sample. This paper presents the first detailed study of the clustering of the DESI-like LRGs.

The selection of our LRG sample is motivated by and intended to mimic the DESI LRG selection. Since spectroscopic redshifts are not available yet, we compute accurate photometric redshifts (photo- z 's) and their error estimates using imaging in g , r , and z bands from Dark Energy Camera Legacy Survey (DECaLS), which is part of the DESI Legacy Imaging Surveys (Dey et al. 2019), and imaging in $W1$ and $W2$ bands from the Wide-field Infrared Survey Explorer (*WISE*; Wright et al. 2010). We use the photo- z 's to measure the projected correlation functions in five redshift bins over redshift range of $0.4 < z \lesssim 0.9$. We interpret the results in the halo occupation distribution (HOD) framework, and we incorporate photo- z error estimates and their uncertainties in this analysis.

Although one can expect that better constraints on HOD parameters will be obtained with spectroscopic redshifts from the upcoming DESI data, this work nevertheless provides an important test of the DESI LRG target selection by providing an estimate of the large-scale structure bias that we can compare with the expected value, and the results also enable the construction of accurate mock galaxy catalogues for DESI analysis before and in the early stage of the survey. Moreover, we demonstrate that data from imaging surveys alone can provide powerful constraints on parameters of HOD and other galaxy–halo connection models.

This paper is organized as follows. We describe the data and the LRG sample in Section 2. We describe the photometric redshifts in Section 3. We describe the clustering measurements in Section 4 and modelling of the measurements in Section 5. We discuss the results and conclude in Sections 6 and 7.

We make the photo- z 's (computed with more recent DR8 data) publicly available.¹ We describe them in Appendix B.

2 DATA

We use the publicly available imaging catalogues from DECaLS DR7 for both sample selection and photometric redshift estimation. DECaLS is one of the DESI imaging surveys, and it provides imaging in $g/r/z$ bands with median 5σ depth of 24.5/24.0/23.0 for the fiducial DESI galaxy target (galaxy with an exponential disc profile with half-light radius of 0.45 arcsec). The source catalogues are constructed using the software package THE TRACTOR (Lang, Hogg & Mykytyn 2016) for source detection and photometry, and they also include the forced photometry of the unWISE coadded images (Lang 2014; Meisner et al. 2019) in the 3.4 ($W1$) and 4.6 micron ($W2$) bands. Some of the data from DES observations are also processed by TRACTOR and included in DR7, and are used here.

2.1 Sample selection

In the final DESI imaging data set, each object on average is covered by three exposures in each of the three optical bands, and to ensure adequate depth and minimize the impact of cosmic rays, we require that each object have at least two exposures in each optical band. We remove objects contaminated by nearby bright stars by applying masks as described in Section 2.2, and we avoid regions of high stellar density by removing the area within $|b| < 25.0^\circ$ (where b is the Galactic latitude).

To facilitate the process of dividing the footprint for jackknife resampling (see Section 4.3), we divide the footprint into HEALPIX (Gorski et al. 2005) pixels of area $\sim 0.21 \text{ deg}^2$ each (corresponding to $N_{\text{side}} = 128$). The pixels at the survey boundaries are removed, and then small ‘islands’ consisting of fewer than 100 pixels are also removed.

The photometry is corrected for Galactic extinction using the Galactic transmission values in the DR7 catalogue. The LRG sample is selected with the following colour and magnitude cuts:

$$(z - W1) > 0.8 \times (r - z) - 0.6, \quad (1a)$$

$$z < 20.41, \quad (1b)$$

$$r - z > (z - 17.18)/2, \quad (1c)$$

$$r - z > 0.9, \quad (1d)$$

$$(r - z > 1.15) \text{ OR } (g - r > 1.65). \quad (1e)$$

Note that all the magnitudes used in this paper are in the AB system. The cuts are shown in Fig. 1. Equation (1a) acts as a stellar rejection cut, similar to the one presented in Prakash et al. (2015). It is shown in the left-hand panel of Fig. 1. This selection relies on the fact that galaxies with old stellar populations have spectra that exhibit a peak at 1.6 micron (sometimes referred to as the ‘1.6 micron bump’); at higher redshift this bump causes an increased flux in the *WISE* $W1$ (3.4 micron) band. Therefore, we can easily separate stars from redshifted galaxies with the $z - W1$ colour. This cut allows us to achieve less than 1 per cent stellar contamination in our sample, as verified using the much deeper and better-seeing HSC data (Aihara et al. 2018a).

Equation (1b) is an apparent magnitude limit for the sample, which is similar to that expected for DESI LRGs (cf. DESI Collaboration 2016b). Equation (1c) is a ‘sliding’ colour–magnitude cut that imposes a redshift-dependent luminosity threshold on the sample, selecting the most luminous galaxies across the redshift range. Therefore, the sliding cut determines the number density and the shape of the redshift distribution. The magnitude limit and sliding cut are shown in the middle panel of Fig. 1.

This sliding cut is combined with the cuts in equations (1d) and (1e), which are shown in the right-hand panel of Fig. 1, to remove low-redshift ($z \lesssim 0.4$) galaxies and select intrinsically red galaxies. These selection cuts yield roughly uniform comoving number density in the redshift range of $0.4 < z \lesssim 0.8$.

2.2 Bright star masks

Objects near bright stars are likely to have inaccurate flux measurements due to contamination. Such contamination causes many objects to be selected as LRGs although their true fluxes do not satisfy the selection cuts. The inaccuracies in the *WISE* point spread function modelling lead to overestimated fluxes for objects near even moderately bright stars, and this is a significant source of imaging systematics for the LRG sample. Another issue is that extremely bright stars produce imaging artefacts such as ghosts and diffraction spikes that are poorly modelled. Such artefacts in the optical imaging, which is used for source detection, cause spurious sources as well as inaccurate flux measurements. Therefore, we apply masks surrounding the positions of bright stars for both optical and *WISE* imaging when constructing the LRG sample.

¹<https://www.legacysurvey.org/dr8/files/#photometric-redshifts>

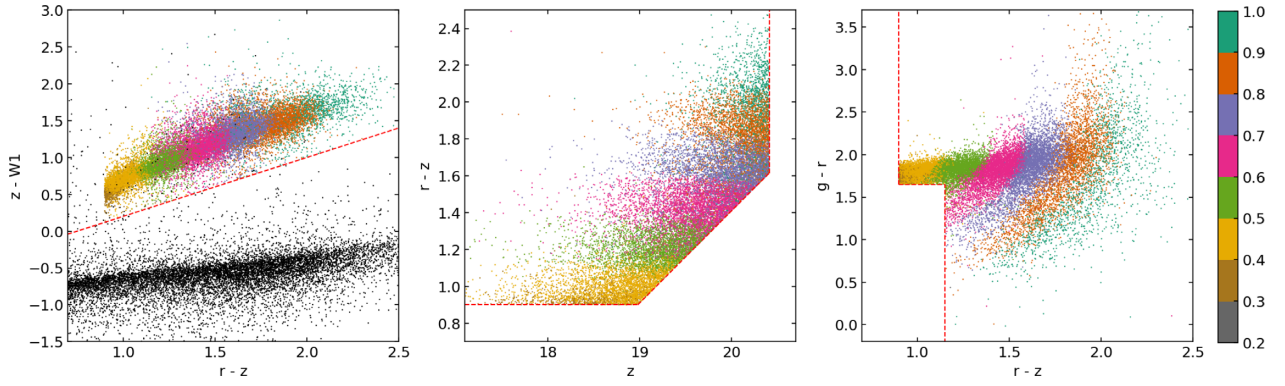


Figure 1. Colour–colour and colour–magnitude diagrams for 20 000 objects randomly selected from the LRG sample. The points are colour coded according to their photometric redshifts (see Section 3). The dashed lines represent the selection boundaries listed in equations (1a)–(1e). The first panel shows the stellar rejection cut using $z - W1$ colour; in this panel, we also plot point sources with $z_{\text{mag}} < 20.41$, the majority of which are stars, in black to show the stellar locus; the stellar rejection cut removes stars very effectively. The second panel shows the sliding colour–magnitude cut and the z -band magnitude cut. The third panels show the cuts that remove low-redshift ($z \lesssim 0.4$) objects.

Three different sets of masks are used. First, we use the ‘bright-star-in-blob’ column in the catalogue. As defined in the DR7 catalogue, a blob is a ‘contiguous region of pixels above a detection threshold and neighbouring pixels’,² and an object is flagged if it is in the same blob as a Tycho-2 star (Høg et al. 2000). Second, we use the unWISE masks described in the appendix of Meisner et al. (2019) to remove areas around bright stars selected from AllWISE (Cutri et al. 2013) and 2MASS (Skrutskie et al. 2006). Third, we develop and apply a set of *WISE* masks that include fainter AllWISE stars that are not in the unWISE masks but still cause significant contamination. The third set of masks is specifically optimized for the LRGs. More about the *WISE* masks can be found in Appendix A. The three sets of masks combined remove ~ 12 per cent of the objects from the LRG sample but only ~ 4 per cent of the total area; it is clear that most of the sources masked do not truly belong in the sample.

In addition to the bright star masks, we also remove regions that are affected by very bright stars or other imaging artefacts. We identify such regions by examining areas with a high density of LRGs that have very large (>0.05) estimated photo- z errors. Since real LRGs typically have much smaller photo- z errors, these objects are mostly either spurious sources or real sources that are not included in the photo- z training such as stars and quasars; however, only the spurious sources are likely to be highly concentrated on the sky (e.g. around very bright stars). Such concentrations are identified efficiently with the DBSCAN cluster analysis routine in SCIKIT-LEARN (Pedregosa et al. 2011) and the corresponding HEALPIX pixels are flagged as bad regions.

The final LRG sample has 2.74 million objects spanning 5655 deg². Fig. 2 shows the sky distribution.

2.3 Randoms

The calculation of correlation functions requires uniformly distributed random points with the same survey geometry as the LRG sample. We use the publicly available random catalogue for DECaLS DR7.³ The same number of exposure requirements, footprint cuts, and bright star masks are applied on the randoms as are used in constructing the LRG sample.

3 PHOTOMETRIC REDSHIFTS

We compute photometric redshifts using the random forest regression method (Breiman 2001), a machine learning (ML) algorithm based on decision trees. For our data set, the ML methods have several advantages over template-fitting methods: there are abundant spectroscopic observations of galaxies covering the magnitude and colour space of the LRG sample that can be used for training, and in this regime ML methods consistently out-perform template-fitting methods; ML methods can trivially incorporate non-photometry information such as galaxy shapes, which we exploit here; and ML methods do not require physical and representative spectral energy distribution (SED) templates, which are not trivial to obtain especially for the wavelength range of the *WISE* passbands. Among the numerous ML methods, random forest provides good performance and is very computationally efficient, so we use it here.

3.1 Imaging data

We include r -band magnitude as well as $g - r$, $r - z$, $z - W1$, and $W1 - W2$ colours as inputs. The photometry has been corrected for Galactic extinction. Soo et al. (2018) showed that while morphological information only provides mild improvements in photo- z accuracy when full *ugriz* photometry is available, the improvement is substantial when only *grz* photometry is available. Motivated by that result, we include as inputs three morphological parameters: half-light radius, axial ratio (ratio between semiminor and semimajor axes), and a ‘model weight’ that characterizes whether a galaxy is better fit by an exponential profile or a de Vaucouleurs, similar to the definition in Soo et al. (2018):

$$p = \frac{d\chi_{\text{dev}}^2}{d\chi_{\text{dev}}^2 + d\chi_{\text{exp}}^2}; \quad (2)$$

where $d\chi^2$ is the difference in χ^2 between the model fit and no source. The inclusion of the three morphological parameters significantly reduces the photo- z scatter for bright galaxies: the normalized median absolute deviation (NMAD), σ_{NMAD} , is reduced by 15 per cent for SDSS (not including BOSS or Extended Baryon Oscillation Spectroscopic Survey, eBOSS) compared to photo- z 's computed without using the morphological parameters; for GAMA galaxies the improvement in σ_{NMAD} is 19 per cent and the 10 per cent outlier fraction is also reduced by ~ 30 per cent. The improvement is less

²<https://www.legacysurvey.org/dr7/description/#glossary>

³<https://www.legacysurvey.org/dr7/files/>

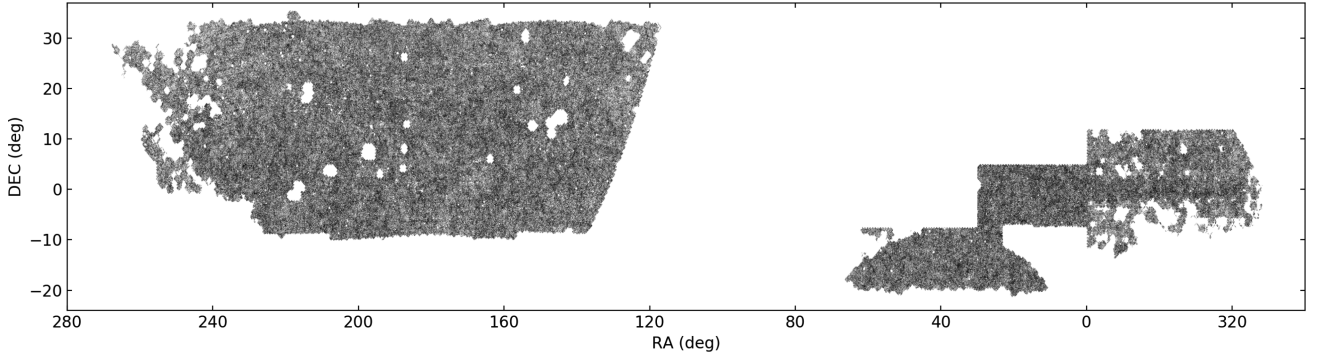


Figure 2. Sky coverage of the final LRG sample that is used in the clustering analysis. The grey-scale represents the surface density. The ‘holes’ in the NGC footprint and parts of the SGC footprint have been removed from the sample due to contamination from very bright stars or other known imaging artefacts.

significant for fainter galaxies. For example, VIPERS galaxies have a 6 per cent improvement in σ_{NMAD} and 12 per cent improvement in the outlier fraction. This is expected, because the morphological parameters of fainter galaxies are less well constrained due to the lower S/N, and because at higher redshifts the angular diameter of a galaxy decreases more slowly with redshift, thus making the galaxy size a less useful quantity for redshift estimation.

3.2 Redshift ‘truth’ data set

For ML photo- z methods, redshift ‘truth’ values are needed for the training process. Various redshift surveys overlap with the DECaLS footprint, and we compile a redshift truth data set using spectroscopic and many-band photometric redshifts from 10 different surveys.

3.2.1 2dFLenS

The 2-degree Field Lensing Survey (Blake et al. 2016) is a spectroscopic survey performed at the Anglo-Australian Telescope, and observed two galaxy samples: a sample of LRGs, and a magnitude limited ($r < 19.5$) galaxy sample. We apply the following quality cuts to this sample:

$$(Q == 4) \text{ AND } (z > 0), \quad (3)$$

where Q is the quality flag.

3.2.2 AGES

The active galactic nucleus (AGN) and Galaxy Evolution Survey (Kochanek et al. 2012) is a spectroscopic survey performed with the Hectospec instrument at the MMT telescope. Targets were selected with optical and infrared imaging down to $I \simeq 20$ (Vega magnitude). Only objects from the galaxy targets are used, and we required $z > 0$.

3.2.3 COSMOS2015 photo- z ’s

The COSMOS2015 catalogue (Laigle et al. 2016) is a photometric redshift catalogue covering the 2 deg² COSMOS field. We include the COSMOS photo- z ’s because there are few spec- z samples with comparable depth and redshift coverage, and adequate coverage in depth and redshift in the training set is necessary for ML-based methods to work. To minimize the risk of photo- z outliers in COSMOS impacting our training, we apply a set of stringent quality cuts motivated by those which were applied in Tanaka et al. (2018):

- (i) FLAG_PETER is false (no bad photometry),
- (ii) TYPE == 0 (only galaxies),
- (iii) $(\text{ZPDF_H68} - \text{ZPDF_L68}) / (1 + z) < 0.02$ (limit photo- z errors to 1 per cent),
- (iv) $(\text{CHI2_BEST} < \text{CHIS})$ AND $(\text{CHI2_BEST} / \text{NBFILT} < 5)$ (fits are reasonable and better than stellar alternatives),
- (v) $\text{ZP_2} < 0$ (no secondary peaks),
- (vi) $\text{MASS_MED} > 7.5$ (stellar mass recovery successful),
- (vii) $\text{DEC} > 1.46$ (removing some apparent imaging artefacts near the lower boundary),
- (viii) $z > 0.006$ (remove the lowest redshift bin),
- (ix) $z < 3$ (redshift upper limit).

We find that if we exclude the COSMOS photo- z ’s from the training sample, the change in photo- z scatter is negligible (< 0.1 per cent of the original value) for spec- z objects with $z_{\text{mag}} < 20.5$; for fainter spec- z objects with $20.5 < z_{\text{mag}} < 21.5$, the photo- z scatter increases by 1 per cent if excluding COSMOS photo- z ’s from the training sample. We also have checked for possible photo- z biases that might be introduced by the COSMOS photo- z ’s by comparing the mean offsets in bins of z_{spec} (e.g. the red line in the lower panel of Fig. B2) with and without COSMOS photo- z ’s in the training sample. We find no noticeable difference in the mean offsets for $z_{\text{mag}} < 20.5$ objects and reduced offsets for $z_{\text{spec}} < 0.5$ objects with $20.5 < z_{\text{mag}} < 21.0$ when COSMOS photo- z ’s are included in the training sample (testing only with objects with fully spectroscopic redshift measurements).

3.2.4 DEEP2

DEEP2 (Newman et al. 2013) is a spectroscopic redshift survey performed with the DEIMOS instrument on the Keck 2 telescope. Targets were selected down to $R_{AB} = 24.1$ with colour cuts to exclude $z < 0.7$ galaxies in three of the four fields surveyed. We require that $z > 0$ and the quality flag $\text{ZQUALITY} \geq 3$.

3.2.5 GAMA DR3

The Galaxy And Mass Assembly (GAMA) survey (Baldry et al. 2018) is a spectroscopic survey performed at the Anglo-Australian Telescope with magnitude limited target selection down to $r \simeq 20$. We require that the quality flag $nQ == 4$ and $z > 0.002$.

3.2.6 OzDES

OzDES is a spectroscopic follow-up survey (Childress et al. 2017) of the DES supernova fields performed at the Anglo-Australian

Telescope. Various types of targets were selected, such as supernova hosts, AGNs, and LRGs. For our purposes, we only use objects that were targeted as ‘LRG’, ‘bright galaxy’, ‘ELG’, ‘photo- z ’, ‘RedMaGiC’, or ‘cluster galaxy’. We also require that the quality flag $Q == 4$ and $z > 0$.

3.2.7 SDSS DR14

We use spectroscopic redshifts from SDSS DR14 (Abolfathi et al. 2018), including the SDSS Main Galaxy Sample (Strauss et al. 2002), BOSS sample (Dawson et al. 2013), and the eBOSS sample (Dawson et al. 2016). To select objects with accurate photo- z 's, we apply the following quality cuts:

- (i) ZWARNING == 0 (no known problems),
- (ii) CLASS == GALAXY (classified as galaxy),
- (iii) $z > 0.0003$ (remove spurious galaxies at very low redshift).

3.2.8 VIPERS

The VIMOS Public Extragalactic Redshift Survey (Scodeggio et al. 2018) is a spectroscopic survey performed at the ESO VLT. The sample is magnitude limited to $i = 22.5$ with colour cuts to exclude $z < 0.5$ galaxies. To ensure redshift quality, we require $3.4 \leq zflg < 5$ where $zflg$ is the quality flag, and $z > 0$.

3.2.9 VVDS

The VIMOS VLT Deep Survey (Le Fevre et al. 2013) is a spectroscopic survey performed at the ESO VLT. The sample is i -band selected, down to $i = 22.5$ in the wide field and $i \simeq 24$ in the deeper field. We require that the quality flag ZFLAGS == 4 and $z > 0$.

3.2.10 WiggleZ

The WiggleZ Dark Energy Survey (Parkinson et al. 2012) is a spectroscopic survey performed at the Anglo-Australian Telescope that aims to measure the BAO signal with emission-line galaxies. We require that the quality flag $Q == 4$ or 5, and we also require $z > 0$.

3.3 Combined truth data set and downsampling

The redshift catalogues are cross-matched to DECaLS with a search radius of 1 arcsec. Fig. 3 shows the redshift distribution of the cross-matched redshift truth data set. Table 1 lists the number of objects from each survey.

Most of the truth objects are from four surveys: SDSS, BOSS, GAMA, and WiggleZ. These surveys either apply specific colour selections (BOSS, WiggleZ) or are limited to shallow magnitudes (SDSS Main Galaxy Sample, GAMA). This leads to sharp peaks in the redshift distribution and discontinuities in the density of objects in colour/magnitude space. Such a non-uniform training sample can cause systematic biases in the photo- z 's; e.g. photo- z algorithms can develop tendency to favour placing objects at redshifts which are overrepresented in training set. To make the training sample more uniform and also to speed up computation, objects from these four large surveys are downsampled. This downsampling is based on the object density in the two-dimensional space of r -band magnitude (used as a proxy for luminosity at fixed redshift) versus redshift, with a bin size of $\Delta z = 0.01$ and $\Delta r_{\text{mag}} = 0.01$. For r_{mag} -redshift bins that have more galaxies than a specific threshold (which are 400, 400,

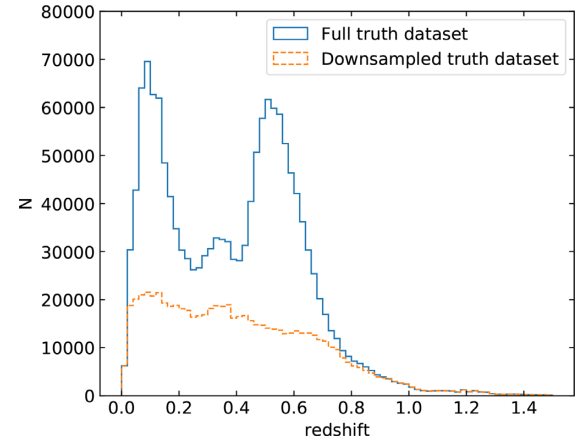


Figure 3. Redshift distribution of the redshift truth data set. $N(z)$ is the total number of objects in each $\Delta z = 0.02$ bin. The peaks at $z = 0.1$ and $z = 0.5$ in the full set are attributed to the SDSS Main Galaxy Sample and BOSS, respectively, both of which are downsampled significantly to avoid biasing the output photo- z 's to favour these redshifts.

Table 1. Number of objects from each redshift survey that are cross-matched to DECaLS.

Survey	Full data set	Downsampled data set
BOSS	678 370	224 345
SDSS	449 386	186 666
WiggleZ	122 907	47 334
GAMA	109 790	55 990
COSMOS2015	53 973	53 972
VIPERS	44 175	44 175
eBOSS	23 549	23 549
DEEP2	15 994	15 994
AGES	11 235	11 235
2dFLenS	8102	8102
VVDS	5490	5490
OzDES	1407	1407

70, and 20 for SDSS, BOSS, GAMA, and WiggleZ, respectively), the objects are randomly downsampled so that the density is reduced to the threshold level. In this way, we reduce the overall number of galaxies while preserving a good sampling of galaxies over the full range of luminosity, in particular retaining the rare most luminous galaxies, many of which are LRGs. The redshift distribution of the downsampled truth data set is shown in Fig. 3. Hereafter, we refer to the downsampled truth catalogue simply as the truth catalogue.

3.4 Random forest method

We compute photo- z 's using the random forest regression routine in SCIKIT-LEARN (Pedregosa et al. 2011). As we have described previously, we use the following eight parameters as input: r -band magnitude, $g - r$, $r - z$, $z - W1$, and $W1 - W2$ colours, half-light radius, axial ratio, and shape probability.

The redshift and magnitude distributions of the photo- z training sample are not uniform due to the various selections of the spectroscopic surveys. In the presence of photometric errors, the gradients in the colour and magnitude distributions cause objects in higher density regions in the multidimensional colour/magnitude space to be scattered into lower density regions. Therefore, in the neighbourhood of each point in the colour/magnitude space, it is more likely to find objects from higher density locations, and since

colours and magnitudes are correlated with redshift (which is why photo- z algorithms work), this causes the photo- z estimates to be biased towards the redshifts of objects in higher density regions in the parameter space. Such bias is particularly significant at the high-redshift end of the redshift distribution, where the photo- z estimates are consistently biased low. To mitigate this bias, we assign weights to each training object based on its spectroscopic redshift (or photometric redshift, in the case of COSMOS): the weight is proportional to the inverse of the number of training objects at that redshift (with a cut-off value to prevent excessively large weights). As a result, objects at very low or very high redshifts are assigned larger weights than other objects. These weights are incorporated into our random forest calculation by requiring a minimum of 25 objects to split an internal node, and then using the weighted average of the redshifts for all training objects in a leaf as its predicted z .

We randomly select 90 percent of the truth data set for training, and reserve the other 10 percent for testing purposes. To estimate the photo- z error for each object, we perturb the photometry by adding to the observed flux in each band a random value from Gaussian distribution whose standard deviation is set by the photometric error. This is similar to Kind & Brunner (2013), although in that work the photometry of the training sample, rather than the sample of objects to which the algorithm is applied as here, was perturbed. Ideally, we would perturb the morphological parameters as well, but were not able to because the errors on these parameters are either not applicable (for the $d\chi^2$ -based ‘model weight’) or not provided (for the half-light radius and the axial ratio) in the imaging catalogue.

For each of the 50 individual trees in the random forest that we generated, we repeat the perturbation 20 times. The mean and standard deviation of the resulting 1000 (50×20) redshift estimates are used as the photo- z and photo- z error, respectively. Note that the photometric noise is only added when computing the final photo- z ’s used for clustering analysis; the random forest is trained with the original unperturbed photometry.

3.5 Photo- z performance for LRGs

Here, we describe the photo- z performance for our LRG sample. For the photo- z performance of the overall spectroscopic training sample, see Appendix B.

To assess the photo- z quality of the LRG sample, we cross-match the LRG sample to the truth catalogues in the multidimensional space of r -band magnitude and $g - r$, $r - z$, $z - W1$, and $W1 - W2$ colours; redshift information is not used in the matching. Each LRG is matched to the single nearest neighbour in the truth catalogue, and we count the number of LRGs that each truth object is matched to. We use this number as the weight for the photo- z versus spec- z plot in Fig. 4. We quantify the photo- z accuracy using the NMAD, defined as $\sigma_{\text{NMAD}} = 1.48 \times \text{median}(|\Delta z|/(1 + z_{\text{spec}}))$ where $\Delta z = z_{\text{phot}} - z_{\text{spec}}$ (we follow the same NMAD definition as in, e.g. Dahlen et al. 2013). This is a robust estimator of scatter as it is not sensitive to outliers. We also measure the fraction of outliers defined as objects with $|\Delta z| > 0.1 \times (1 + z_{\text{spec}})$. From the weighted spec- z objects, we estimate that the average photo- z scatter σ_{NMAD} for the LRG sample is 0.021 and the outlier rate is 1 per cent. The weighted spec- z objects are also used for estimating the comoving number density which we use as an observable in our HOD modelling (see Section 5).

3.5.1 Photo- z error estimates

The photo- z error estimation procedure described in Section 3.4 is built upon the assumptions that (1) the photo- z probability density

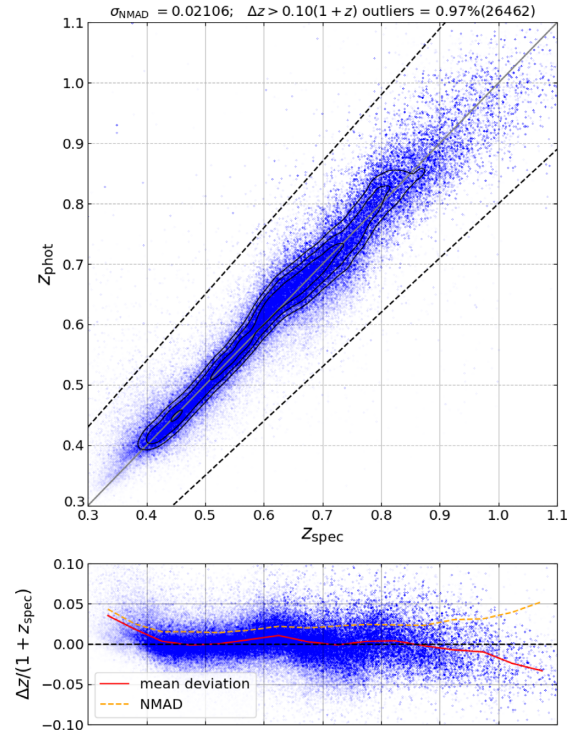


Figure 4. Top panel: Photo- z versus spec- z for truth objects that are weighted to approximate the photo- z performance of the LRG sample. Contours are used to indicate density where the points are densest. Black dashed lines indicate the boundaries outside which objects are classified as outliers. Lower panel: photo- z offset (in $\Delta z/(1 + z_{\text{spec}})$) versus spec- z ; the red solid line and the yellow dashed line are the Hodges–Lehmann mean and σ_{NMAD} , respectively, of the photo- z offset in bins of spec- z . The photo- z ’s are mostly well constrained with few outliers for this sample.

functions (PDFs) are Gaussian; (2) photo- z uncertainties are dominated by photometric errors; and (3) the photometric error estimates are accurate. They do not include the effects of incompleteness in the training data or uncertainties in morphological parameters.

We have validated the photo- z error estimates using objects with spectroscopic redshifts. If the photo- z PDFs are indeed Gaussian and the estimated photo- z errors are accurate, the photo- z of each galaxy should differ from its spec- z by a random value Δz drawn from a Gaussian distribution with σ equal to the estimated photo- z error, i.e. $\Delta z/\sigma_z$, where σ_z is the estimated photo- z error for a given object, should be a standard normal distribution with mean 0, and standard deviation 1. As shown in Fig. 5, we find that $\Delta z/\sigma_z$ is roughly consistent with a Gaussian distribution, but the photo- z errors are generally overestimated by a factor which depends on the redshift range. Because we do not have representative spectroscopic data to estimate this scaling factor accurately, we treat it as nuisance parameter in the HOD modelling (see Section 5.2).

4 CLUSTERING MEASUREMENTS

4.1 Redshift bins

In order to study the redshift dependence of the sample properties, we divide the LRG sample into four photometric redshift bins of width $\Delta z \simeq 0.1$, covering the ranges [0.41, 0.5], [0.5, 0.61], [0.61, 0.72], [0.75, 0.83], and [0.84, 0.93]. These bins have been chosen such that they are centred at the redshifts of the snapshots of the halo catalogues

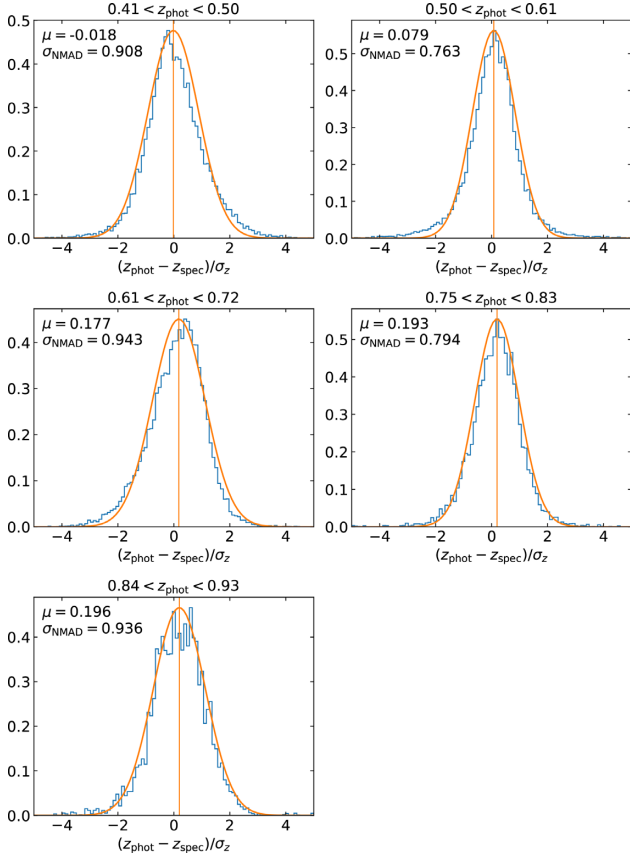


Figure 5. The distribution of $(z_{\text{phot}} - z_{\text{spec}})/\sigma_z$, where σ_z is the estimated photo- z error, in different redshift bins. Here, μ (vertical line) and σ_{NMAD} (1σ width of the curve) are the median and the NMAD of the distribution, respectively, and the smooth curves show the corresponding Gaussian distributions. A non-zero μ value indicates that the photo- z 's are biased, and any deviation of σ_{NMAD} from unity indicates over or underestimation of the photo- z errors. The fact that σ_{NMAD} values are consistently less than unity indicates that our photo- z errors are overestimated.

from the N -body simulation which we use for the clustering analysis in Section 5 (there are small gaps between some of the redshift bins for the same reason). Fig. 6 shows the estimated comoving number density versus redshift for the LRG sample; the shaded regions represent the redshift ranges for each bin. The volume-averaged comoving number densities for each bin are listed in Table 2. The densities estimated from photo- z 's are consistent with the estimates from weighted spec- z 's. Densities estimated using spec- z 's weighted as described in Section 3.5 are used in modelling.

Fig. 7 shows the redshift distribution (surface density) for each redshift bin estimated by convolving the photo- z 's with their photo- z error.

4.2 Projected correlation function

Ideally, one would like to measure the correlation function $\xi(r)$, the excess probability of finding a pair of galaxies separated by distance r , but the large uncertainty in radial distances precludes its direct measurement. For imaging data sets, it is common to measure instead the angular correlation function, typically in bins of photometric redshifts. However, in doing so the information on relative distances between galaxies contained in the photo- z 's is not utilized. For spectroscopic data sets, due to the presence of redshift-

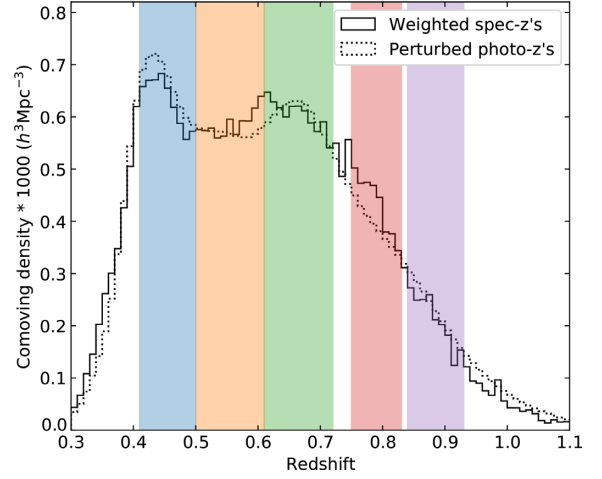


Figure 6. Redshift-dependent comoving number density of the LRG sample. The dashed line shows the densities estimated with photo- z 's convolved by the estimated photo- z error. The solid line shows the densities estimated from the weighted spec- z 's described in Section 3.5; these weighted spec- z 's are also used for deriving the comoving number densities for the clustering analysis. The coloured bands represent the redshift bins. The fact that the two lines are consistent with each other indicates that our comoving density estimates are relatively robust. The uncertainties in the densities are accounted for in our modelling (see Section 5.3).

Table 2. The redshift bins. The second column lists the redshifts of the snapshots of the N -body simulation. The third column lists the comoving number densities in units of $h^3 \text{ Mpc}^{-3}$; these values are used in the HOD fitting with 10 per cent assumed Gaussian uncertainty.

Redshift	z_{sim}	$n(z)$
$0.41 < z_{\text{phot}} < 0.50$	0.4573	6.32×10^{-4}
$0.50 < z_{\text{phot}} < 0.61$	0.5574	6.16×10^{-4}
$0.61 < z_{\text{phot}} < 0.72$	0.6644	6.15×10^{-4}
$0.75 < z_{\text{phot}} < 0.83$	0.7787	4.41×10^{-4}
$0.84 < z_{\text{phot}} < 0.93$	0.8594	2.14×10^{-4}

space distortions, the small-scale clustering is usually measured with the projected correlation function, effectively integrating out the effects of redshift-space distortions:

$$w_p(r_p) = \int_{-\pi_{\text{max}}}^{\pi_{\text{max}}} \xi(r_p, \pi) d\pi, \quad (4)$$

where $\xi(r_p, \pi)$ is the 3D correlation function, r_p is the transverse distance, and π is the line-of-sight distance.

Here, to better exploit the photo- z information, we measure the projected correlation function (equation 4), using distances derived from the photo- z 's. Fig. 8 illustrates our method of measuring $w_p(r_p)$. We adopt a relatively large π_{max} of $150 h^{-1} \text{ Mpc}$ to account for the large radial distance uncertainties from the photo- z 's. This large π_{max} is comparable to the width of the redshift bins ($160\text{--}210 h^{-1} \text{ Mpc}$). However, the photo- z errors cause many galaxy pairs to be lost due to one of the galaxies being outside of the redshift bin in a generic autocorrelation measurement, thus complicating the modelling and resulting in a lower clustering signal-to-noise ratio. To address this issue, instead of counting pairs within the redshift bin z_i (i.e. measuring a simple autocorrelation), we define a wider redshift bin $z_{\text{wide}, i}$ that encloses z_i , and count pairs between galaxies in z_i and galaxies in $z_{\text{wide}, i}$. The wider redshift bin $z_{\text{wide}, i}$ extends from z_i

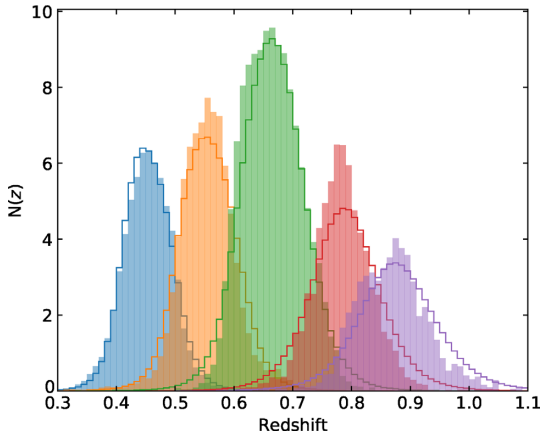


Figure 7. The estimated redshift distributions (surface density in redshift bins) of the LRG sample. The y-axis is the number of objects per sq. degree in the redshift bin of width 0.1. The filled histograms show the redshift distributions estimated from weighted spec- z objects. The unfilled histograms are distributions of stacked photo- z 's that are convolved with photo- z errors (however, objects are assigned to a given redshift bin based only upon their random forest point estimates). The samples in different redshift (photo- z) bins overlap due to photo- z errors. Note that (1) we use the original photo- z error estimates, which as seen in Fig. 5, are probably overestimated; and (2) this simple stacking of the photo- z probabilities produces broader redshift distributions than the mathematically correct procedure (see Malz & Hogg 2020).

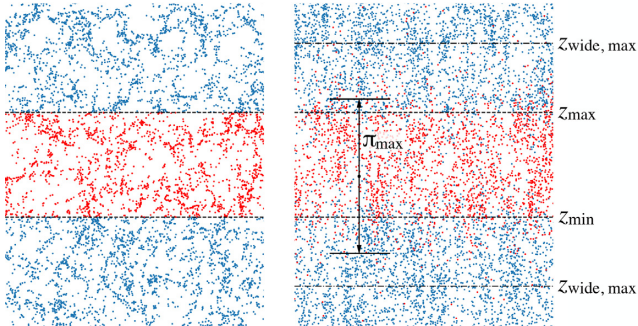


Figure 8. Illustration of our method for measuring $w_p(r_p)$ with photo- z 's; differences in the redshift direction are shown to scale for the $0.61 < z_{\text{phot}} < 0.72$ redshift bin. *Left-hand panel:* galaxies with accurate redshifts. The dashed lines show the redshift bin boundary, and the galaxies inside and outside the redshift bin are colour coded in red and blue, respectively. *Right-hand panel:* the same galaxies with the same colour coding as the left-hand panel, but with realistic redshift (photo- z) uncertainties added. A significant number of galaxies cross the redshift boundaries. The arrows show the $\pm\pi_{\text{max}}$ range with $\pi_{\text{max}} = 150 h^{-1}$ Mpc used in counting pairs and calculating $w_p(r_p)$. The dash-dotted line shows the wider redshift bin used in the 'padded' $w_p(r_p)$ calculation; by construction it includes all objects that are within $\pm\pi_{\text{max}}$ of any object with $0.61 < z_{\text{phot}} < 0.72$.

by π_{max} in both directions, so that for all galaxies in $z_{\text{wide}, i}$, all their pairs within π_{max} will be counted. Besides boosting the clustering S/N, this 'padded' autocorrelation approach decouples the clustering measurement from any effects associated with the boundary of the redshift bin, and thus simplifies the measurement and modelling of the projected correlation function.

To compute this 'padded' $w_p(r_p)$, we adopt the cross-correlation form of the Landy-Szalay estimator (Landy & Szalay 1993):

$$w_p(r_p) = 2\pi_{\text{max}} \times \frac{(D_1 D_2 - D_1 R_2 - D_2 R_1 + R_1 R_2)}{R_1 R_2}, \quad (5)$$

where each term denotes the pair count between two samples; D_1 denotes the data (galaxy sample) in a redshift bin defined in Section 4.1 and D_2 denotes the data in the wider redshift bin; and R_1 and R_2 denote sets of random points with the same angular and redshift distribution as D_1 and D_2 , respectively. The redshifts of the randoms are randomly drawn from the redshifts of the data, so they have the same redshift distribution as the data, by construction. In each redshift bin, the ratio of the number of randoms to data is 20. The measurement is done using the TREECORR package (Jarvis, Bernstein & Jain 2004). We tested this estimator on mocks that resemble our redshift bins, and we confirmed that it produces the same $w_p(r_p)$ as that for mocks in a cubic volume with periodic boundary conditions using the simple autocorrelation estimator.

We measure $w_p(r_p)$ in 12 logarithmically spaced bins covering the range from 0.11 to $19.5 h^{-1}$ Mpc (comoving). We also made measurements at smaller and larger scales, but these measurements were not used for the modelling. Fig. 9 shows the measured projected correlation functions for the five redshift bins. The two bumps, corresponding to the one-halo (from galaxy pairs within the same dark matter halo) and two-halo (from galaxy pairs in two different haloes) regimes of the correlation function, are clearly visible in all the redshift bins. It should be noted that the photo- z errors effectively smooth out sharp features in the radial direction, and as a result these $w_p(r_p)$ measurements using photo- z 's have lower amplitude than the intrinsic clustering signal that one can measure with spectroscopic redshifts.

As shown in Fig. 7, the sample in each redshift bin does not contain all the objects in that redshift range, and it also includes objects whose true redshift is in other bins.

4.3 Jackknife resampling and covariances

We compute the covariance matrices of correlation functions with jackknife resampling: we divide the footprint into N_{sub} subregions, and we resample the data set with one of the subregions removed. There are a total of N_{sub} resampled data sets, and the correlation function is measured for each one. The covariance matrix is given by

$$\text{Cov}(w_i, w_j) = \frac{(N_{\text{sub}} - 1)}{N_{\text{sub}}} \sum_{l=1}^{N_{\text{sub}}} (w_i^l - \bar{w}_i) (w_j^l - \bar{w}_j), \quad (6)$$

where w_i^l is the projected correlation function at the i -th distance bin measured from the l -th jackknife sample, and \bar{w}_i is the mean from all jackknife samples.

The DECaLS survey was not yet completed by DR7. This, combined with the masks and quality cuts, results in the irregular survey geometry, making it difficult to manually divide the footprint into compact subregions with equal areas. To solve this problem, we developed an automated routine. The relevant PYTHON codes can be found online.⁴ The routine involves three steps:

- (i) The objects (in our case randoms) in the survey footprint are divided into HEALPIX pixels. This significantly reduces the number of points and speeds up computation;
- (ii) Initial grouping of the HEALPIX pixels is performed using a clustering algorithm (specifically, we use k-means clustering; see Hartigan & Wong 1979) with the object count in each pixel as weights to account for fractional occupation of the pixels;

⁴https://github.com/rongpu/pixel_partition

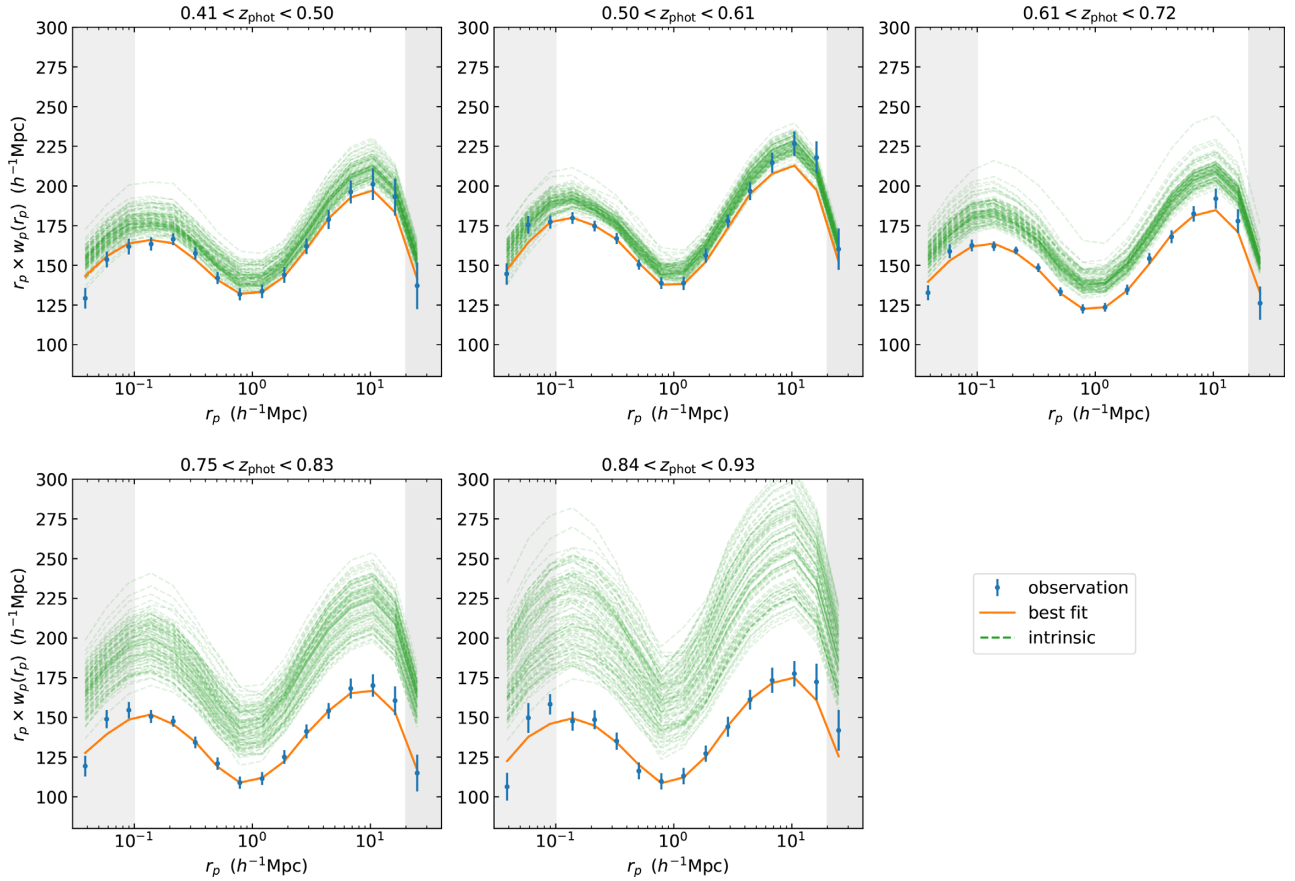


Figure 9. The projected correlation function multiplied by the transverse distance. The points are measurements with error bars from jackknife. The points in the grey shaded area are not used for modelling to avoid possible systematics. The orange curve is the best fit from HOD modelling. The green band is the [16th, 84th] percentile range of the intrinsic clustering signal, i.e. the clustering that would have been measured according to the fit parameters if perfect distance measurements (rather than photo- z 's) were available. The larger uncertainty in the intrinsic clustering in the higher redshift bins is mostly due to the larger photo- z errors for them.

(iii) We change the group labels of a specific fraction of boundary pixels, which were randomly selected, to the labels of their neighbouring group(s), and only keep these changes if they improve the ‘score’ which is a weighted sum of the compactness (defined as the average of the standard deviations of the angular distances, in arcseconds, between the positions of all objects in a subregion and its centre) and uniformity of the areas (defined as the standard deviation of the areas of the subregions);

(iv) We repeat this step until the desired score is achieved.

Specifically, in each iteration the fraction of changed boundary pixels is 0.15 percent; and equal weights were given to uniformity and compactness for calculating the score.

Applying this procedure, we divide the footprint into 120 jackknife subregions of 47.1 deg^2 each (within ~ 2 per cent variation), shown in Fig. 10. The areas of the subregions are uniform to ~ 1 per cent. The area of the subregions is large enough to cover the entirety of angular scales of interest, and small enough to produce enough subregions to compute covariance matrices accurately. The size of each subregion is much larger than the largest angular scale of ~ 1 deg in the correlation function, therefore the jackknife resampling accounts for cosmic variance in addition to shot noise and imaging systematics.

5 MODELLING

In this section, we present our analysis of the galaxy–halo connection for the LRG sample using the HOD framework. The galaxy–halo

connection determines the clustering properties of the galaxy sample, such as the large-scale bias, which is useful for forecasting DESI BAO constraints. Predictions of the locations of galaxies from simulations of dark matter alone also form the basis of mock galaxy catalogues that can be used for estimating covariances between measured large-scale structure statistics. In this work, we fix the cosmological parameters and only allow the HOD parameters to vary, but in principle the method presented here can be incorporated into a more flexible modelling framework that constrains both cosmology and galaxy–halo connection parameters.

5.1 HOD model

We fit the measured clustering signal with an HOD model (see e.g. Berlind & Weinberg 2002 which also lists earlier literature on HOD; see also Wechsler & Tinker 2018 for a more general review on the galaxy–halo connection), which is widely used to model luminosity-threshold galaxy samples. In this framework, dark matter haloes from N -body simulations are populated by central and satellite galaxies with a probabilistic prescription. In its basic form which we have adopted, the central galaxy probability (denoted by N_{cen}) and the mean number of satellite galaxies (denoted by N_{sat}) for a given a halo are determined solely by the halo mass. There are several slightly different mathematical prescriptions for this. We adopt the one in Zentner et al. (2019); we briefly summarize it below.

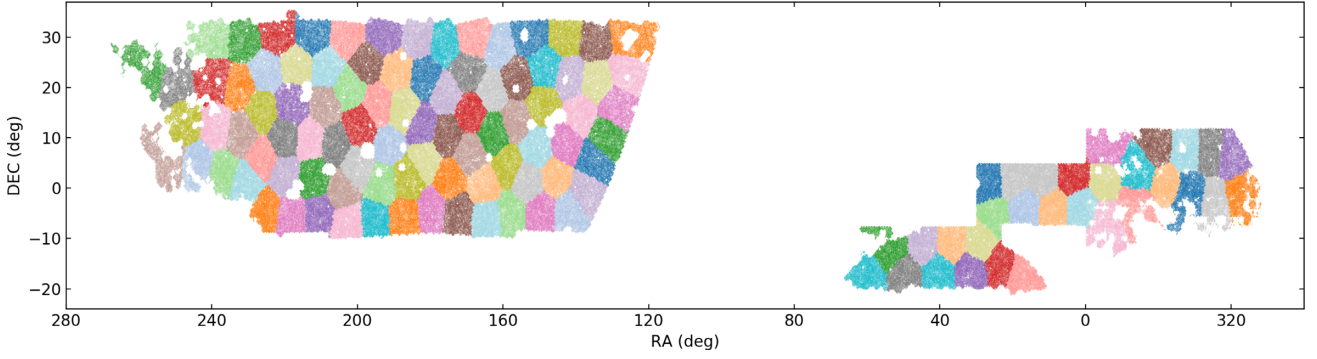


Figure 10. Sky distribution of the LRGs with colour coding to show the 120 subregions for jackknife resampling. Note that non-adjacent ‘patches’ with the same colour are different subregions. Each subregion has the same area (within ~ 2 per cent variation) and is compact by design.

In this formulation, the central galaxy probability is given by a step-like function

$$\langle N_{\text{cen}} | M_{\text{vir}} \rangle = \frac{1}{2} \left(1 + \text{erf} \left[\frac{\log(M_{\text{vir}}) - \log(M_{\text{min}})}{\sigma_{\log M}} \right] \right), \quad (7)$$

where M_{vir} is the virial mass of the dark matter halo; M_{min} is the mass threshold above which haloes are populated by central galaxies; and $\sigma_{\log M}$ defines the smoothness of this transition.

The number of satellite galaxies in a halo follows the Poisson distribution with a mean given by a power law

$$\langle N_{\text{sat}} | M_{\text{vir}} \rangle = \left(\frac{M_{\text{vir}} - M_0}{M_1} \right)^\alpha, \quad (8)$$

where M_0 , M_1 , and α are free parameters of the HOD model. We impose that there are no satellite galaxies in haloes with $M_{\text{vir}} < M_0$. We allow satellite galaxies to exist in haloes with no central galaxies.

The spatial distribution of satellite galaxies is assumed to follow the NFW profile (Navarro, Frenk & White 1997), with the scale radius R_s given by the value for the parent halo in the N -body catalogue. We ignore the effect of velocity dispersion on redshift since it is negligible compared to photo- z errors.

5.2 Mock galaxies

To constrain the HOD model parameters, we measure the clustering of mock galaxies generated from a set of assumed values for these parameters and compare with our measurements from real LRGs. We use the halo catalogue from the MultiDark Planck 2 (MDPL2) simulation (Klypin et al. 2016), which used ROCKSTAR (Behroozi, Wechsler & Wu 2013) to identify the haloes. The MDPL2 simulation adopts the Planck 2013 cosmology (Planck Collaboration XVI 2014): $\Omega_m = 0.307115$, $\Omega_\Lambda = 1 - \Omega_m = 0.692885$, $\Omega_b = 0.048206$, $h = 0.6777$, $\sigma_8 = 0.823$, and $n_s = 0.96$. The size of the cubic simulation box is $1 h^{-1}$ Gpc, and the mass resolution is $1.51 \times 10^9 h^{-1} M_\odot$. We designed the redshift bins of the LRG sample so that they centre at the redshifts of five snapshots of the simulation, as listed in Table 2. We use HALOTOOLS (Hearin et al. 2017) for populating the haloes with galaxies using the pre-defined Zheng, Coil & Zehavi (2007) prescription (note that we use the default HALOTOOLS definition of $\langle N_{\text{sat}} | M_{\text{vir}} \rangle$ which is slightly different from Zheng et al. 2007).

To emulate the effect of photo- z ’s on the clustering signal, we perturb the position of the galaxies along the line-of-sight direction, i.e. the direction along one of the axes of the simulation box. For each galaxy, this distance perturbation is drawn from a Gaussian distribution, the width of which is randomly drawn from the rescaled photo- z error estimates of the LRGs in the corresponding redshift bin.

Table 3. The ranges of the flat priors on model parameters.

Parameter	Prior interval
$\log(M_{\text{min}})$	[11.0, 14.0]
$\sigma_{\log M}$	[0.001, 1.5]
α	[0.0, 2.0]
$\log(M_0)$	[11.0, 14.0]
$\log(M_1)$	[11.5, 15.5]
S_z	[0.6, 1.4]

As discussed in Section 3.5.1, the true photo- z errors are assumed to differ from the estimated errors by a scaling factor S_z which we do not have good constraints on. Thus S_z is included in the model as a nuisance parameter. We use the CORRFUNC software package (Sinha & Garrison 2017) for measuring $w_p(r_p)$ of the mock galaxies because it is better optimized than TREECORR for cubic boxes with periodic boundaries conditions.

To test whether allowing skewness in the photo- z PDF would change our results, we have repeated the analysis using a skew normal distribution as the photo- z PDF for one of the redshift bins, with skewness derived by fitting the $(z_{\text{phot}} - z_{\text{spec}})/\sigma_z$ distribution. We find negligible changes in the results; see Appendix C for more details.

5.3 MCMC sampling of parameters

As described in Section 5.1, the HOD model used here has five free parameters: $\log(M_{\text{min}})$, $\sigma_{\log M}$, α , $\log(M_0)$, and $\log(M_1)$. Additionally, we have the nuisance parameter S_z to account for uncertainties in the photo- z error estimation. We adopt flat priors for all these parameters, and the ranges for the priors are listed in Table 3. We set the lower limit of S_z at 0.6, which is much lower than the values inferred from spectroscopic redshifts (see Fig. 5).

To obtain the posterior probability distributions of the HOD parameters, we perform Markov chain Monte Carlo (MCMC) sampling using the EMCEE package (Foreman-Mackey et al. 2013). The likelihood function used is given by $\mathcal{L} \propto e^{-\chi^2/2}$, where χ^2 is given by

$$\chi^2 = \Delta w_i [\text{Cov}^{-1}]_{ij} \Delta w_j + \frac{(n^{\text{meas}} - n^{\text{mock}})^2}{\sigma_n^2}, \quad (9)$$

where $\Delta w_i = w_p^{\text{meas}}(r_p) - w_p^{\text{mock}}(r_p)$ is the difference between the measured projected correlation function and the one from the mocks at the i -th distance bin; Cov^{-1} is the inverse of the covariance matrix from jackknife resampling (see Section 4.3); n^{meas} and

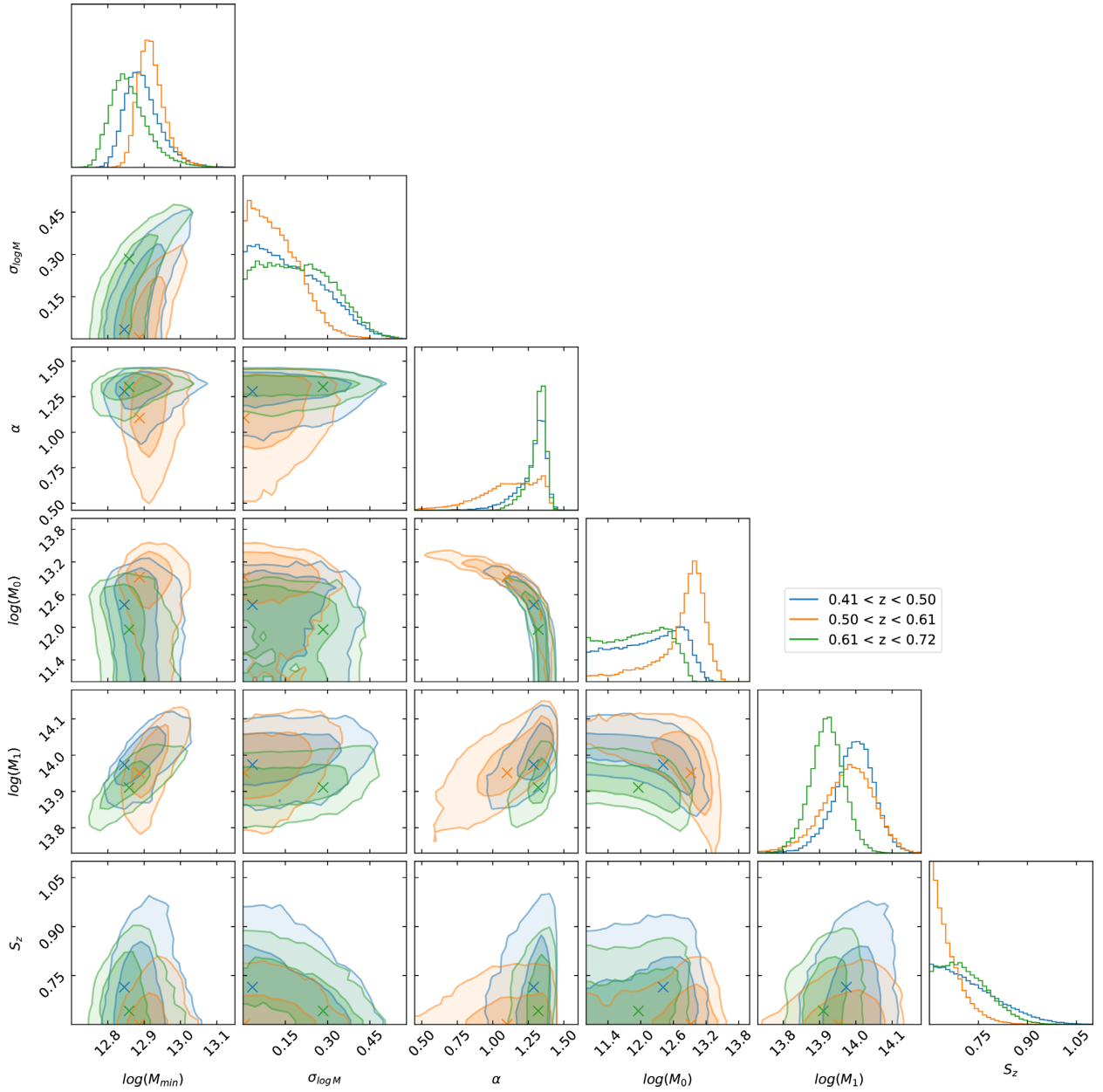


Figure 11. The one- and two-dimensional marginalized posterior probability distributions of the model parameters from MCMC for the first three redshift bins. The inner contour and outer contour are the 68 per cent and 95 per cent confidence regions, respectively. All parameters except S_z are HOD parameters; the parameter S_z is the scaling factor for the photo- z errors. The mass parameters are in units of $h^{-1} M_\odot$.

n^{mock} are the comoving number densities of the data and the mock, respectively; and σ_n is the uncertainty in the comoving number densities. The comoving number densities are estimated from the weighted spec- z objects as described in Section 3.5. We have compared these density estimates with the densities estimated from the photo- z 's, and the standard deviation of the differences in the five redshift bins is ~ 7 per cent. We adopt a larger density uncertainty of $\sigma_n = 10$ per cent to account for any additional unknown systematics in the photo- z 's.

6 RESULTS

The posterior distributions of the model parameters are shown in Figs 11 and 12. These plots have been made using a modified version

of the CORNER.PY software package (Foreman-Mackey 2016). The mean, median, and best-fitting values of the parameters and the 16 per cent and 84 per cent percentiles are listed in Table 4. The derivation of the best-fitting parameters is described in Appendix D. The first four redshift bins have very similar HOD parameters, indicating that the LRGs have similar host halo properties over the redshift range of $0.4 < z < 0.8$.

The highest-redshift sample has much wider contours on the model parameters than lower redshift samples. This could be due to a combination of several factors, including the larger photo- z errors at higher redshifts, a smaller sample size, and a broader redshift distribution that dilutes the clustering signal. Despite the larger errors, the differences in the model parameters between this and lower redshift samples are statistically significant. Such deviation is caused by the selection effect of the apparent magnitude limit

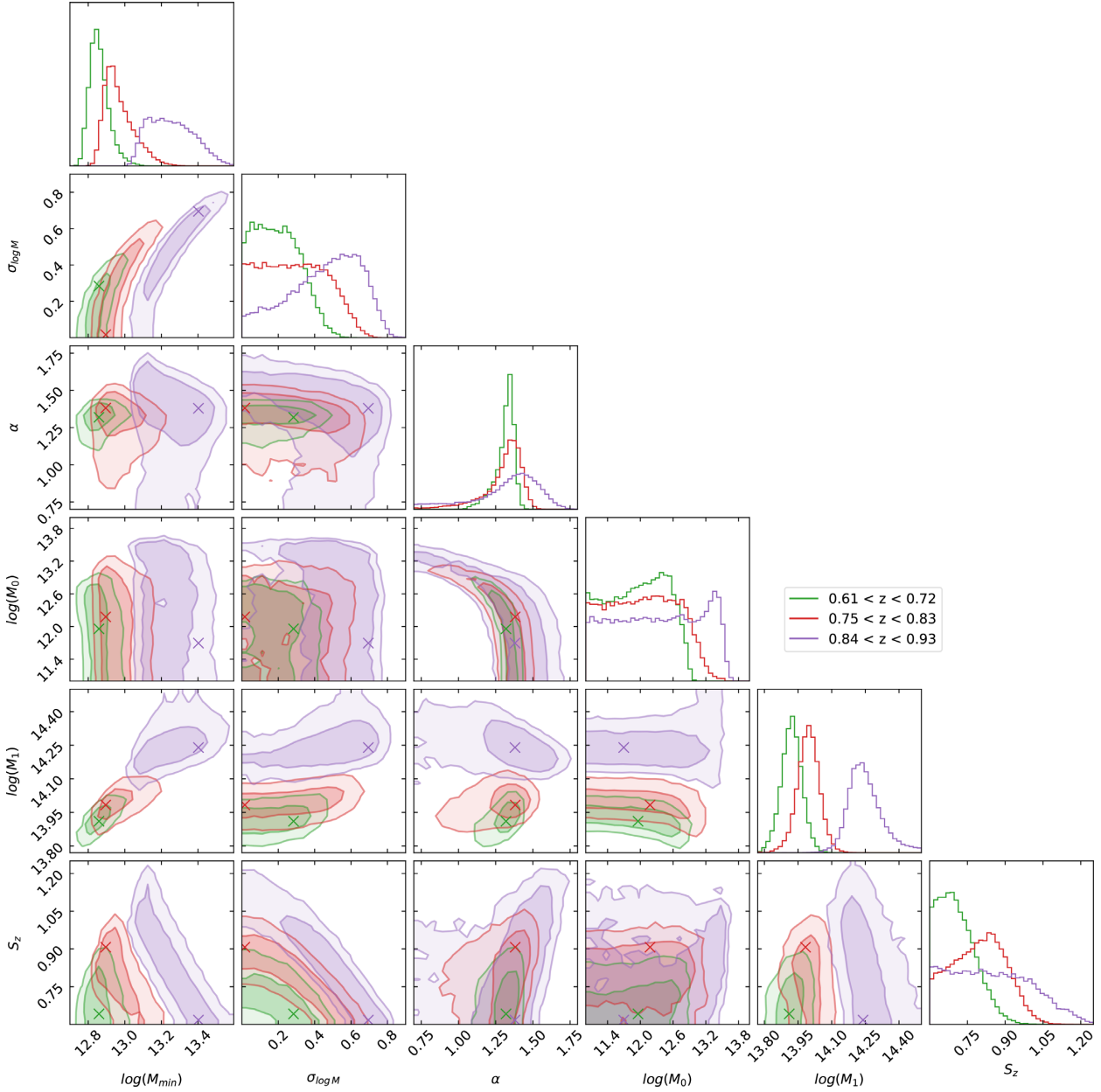


Figure 12. Same as Fig. 11, but for different redshift bins. We include the results of $0.61 < z_{\text{phot}} < 0.72$ (which was also shown in the previous plot). The contours of the highest-redshift bin ($0.84 < z_{\text{phot}} < 0.93$) are significantly different from those of the other redshift bins.

(see the middle panel of Fig. 1): at redshifts higher than ~ 0.7 , the luminosity threshold established by the sliding cut is replaced by the apparent magnitude limit, and as a result more luminous galaxies are more and more preferentially selected as redshift increases. This qualitative change of the sample at high redshift will need to be considered when analysing and interpreting DESI spectroscopic data.

At lower redshifts where the photo- z errors are small, the HOD parameters are not sensitive to the photo- z error rescaling factor S_z , since most of the galaxy pairs are still within $\pi_{\text{max}} = 150 \, h^{-1} \text{ Mpc}$ in the line-of-sight direction. At higher redshifts, the photo- z errors are much larger, and the clustering signal from the mock galaxies is much more sensitive to the S_z , thus resulting in the strong correlation between some of the HOD parameters and S_z . Nevertheless, the value of S_z is poorly constrained by the data, although the posterior prefers $S_z \lesssim 1$ at all redshifts, which is consistent

with the results of the photo- z error validation with spectroscopic redshifts.

Fig. 13 shows the halo occupation functions. Fig. 14 shows the probability distributions for the host halo masses in each redshift bin. In both figures, the solid lines are the best-fitting results, and the dashed lines are for parameters randomly selected from the MCMC chains to show the possible range of halo occupations allowed by the data. The particularly sharp transition (quantified by the $\sigma_{\log M}$ parameter) in the central galaxy population might be due to insufficient flexibility in the HOD model or inaccuracy in the σ_8 parameter we assumed, and the strong correlation between $\sigma_{\log M}$ and the photo- z error rescaling factor also suggests the possibility of biased $\sigma_{\log M}$ estimates due to additional photo- z systematics that is not accounted for by the rescaling factor.

Ideally, we would compute the galaxy bias by comparing the clustering amplitudes of galaxies and dark matter particles in the

Table 4. Results from the HOD fitting with MCMC for the five redshift bins. For each redshift bin, the first row lists the mean values and the 16th and 84th percentiles; the second row (in *italics* font) lists the median values; and the third row (in **bold** font) lists the best-fitting values in the six-dimensional HOD parameter space (these need not match the peak of each marginalized posterior distribution). The parameters $\log(M_{\min})$, $\sigma_{\log M}$, α , $\log(M_0)$, and $\log(M_1)$ are free parameters. The bias b and satellite fraction f_{sat} are derived parameters. We also list the χ^2 for the best-fitting parameters and the corresponding one-sided p -value; the fit quality is acceptable for all five redshift bins (values of $p < 0.05$ would indicate statistically significant differences between the model and the data).

Redshift	$\log(M_{\min})$	$\sigma_{\log M}$	α	$\log(M_0)$	$\log(M_1)$	b	f_{sat}	χ^2 (p -value)
$0.41 < z_{\text{phot}} < 0.50$	$12.90^{+0.05}_{-0.05}$	$0.17^{+0.13}_{-0.13}$	$1.28^{+0.09}_{-0.10}$	$12.13^{+0.64}_{-0.71}$	$13.99^{+0.05}_{-0.05}$	$1.89^{+0.03}_{-0.03}$	$0.15^{+0.02}_{-0.02}$	4.57 (0.71)
	<i>12.89</i>	<i>0.16</i>	<i>1.31</i>	<i>12.19</i>	<i>14.00</i>	<i>1.88</i>	<i>0.15</i>	
	12.85	0.03	1.29	12.41	13.97	1.89	0.15	
$0.50 < z_{\text{phot}} < 0.61$	$12.92^{+0.03}_{-0.04}$	$0.12^{+0.09}_{-0.09}$	$1.11^{+0.22}_{-0.21}$	$12.73^{+0.40}_{-0.46}$	$13.98^{+0.07}_{-0.07}$	$2.04^{+0.02}_{-0.02}$	$0.13^{+0.01}_{-0.01}$	9.87 (0.20)
	<i>12.92</i>	<i>0.11</i>	<i>1.14</i>	<i>12.88</i>	<i>13.98</i>	<i>2.04</i>	<i>0.13</i>	
	12.89	0.00	1.10	12.92	13.95	2.04	0.13	
$0.61 < z_{\text{phot}} < 0.72$	$12.86^{+0.05}_{-0.05}$	$0.20^{+0.13}_{-0.13}$	$1.31^{+0.06}_{-0.06}$	$11.94^{+0.58}_{-0.62}$	$13.92^{+0.04}_{-0.04}$	$2.07^{+0.03}_{-0.03}$	$0.16^{+0.01}_{-0.01}$	6.86 (0.44)
	<i>12.85</i>	<i>0.19</i>	<i>1.33</i>	<i>11.97</i>	<i>13.92</i>	<i>2.07</i>	<i>0.16</i>	
	12.86	0.28	1.32	11.96	13.91	2.04	0.16	
$0.75 < z_{\text{phot}} < 0.83$	$12.97^{+0.08}_{-0.08}$	$0.29^{+0.19}_{-0.20}$	$1.28^{+0.12}_{-0.12}$	$12.05^{+0.69}_{-0.71}$	$14.00^{+0.05}_{-0.05}$	$2.24^{+0.07}_{-0.08}$	$0.14^{+0.02}_{-0.02}$	3.29 (0.86)
	<i>12.95</i>	<i>0.28</i>	<i>1.33</i>	<i>12.06</i>	<i>14.00</i>	<i>2.24</i>	<i>0.14</i>	
	12.90	0.02	1.38	12.18	13.98	2.31	0.14	
$0.84 < z_{\text{phot}} < 0.93$	$13.26^{+0.13}_{-0.13}$	$0.45^{+0.20}_{-0.22}$	$1.20^{+0.31}_{-0.42}$	$12.33^{+0.91}_{-0.90}$	$14.31^{+0.07}_{-0.12}$	$2.48^{+0.15}_{-0.15}$	$0.10^{+0.02}_{-0.02}$	6.65 (0.47)
	<i>13.24</i>	<i>0.48</i>	<i>1.34</i>	<i>12.34</i>	<i>14.25</i>	<i>2.47</i>	<i>0.10</i>	
	13.40	0.70	1.38	11.70	14.24	2.28	0.10	

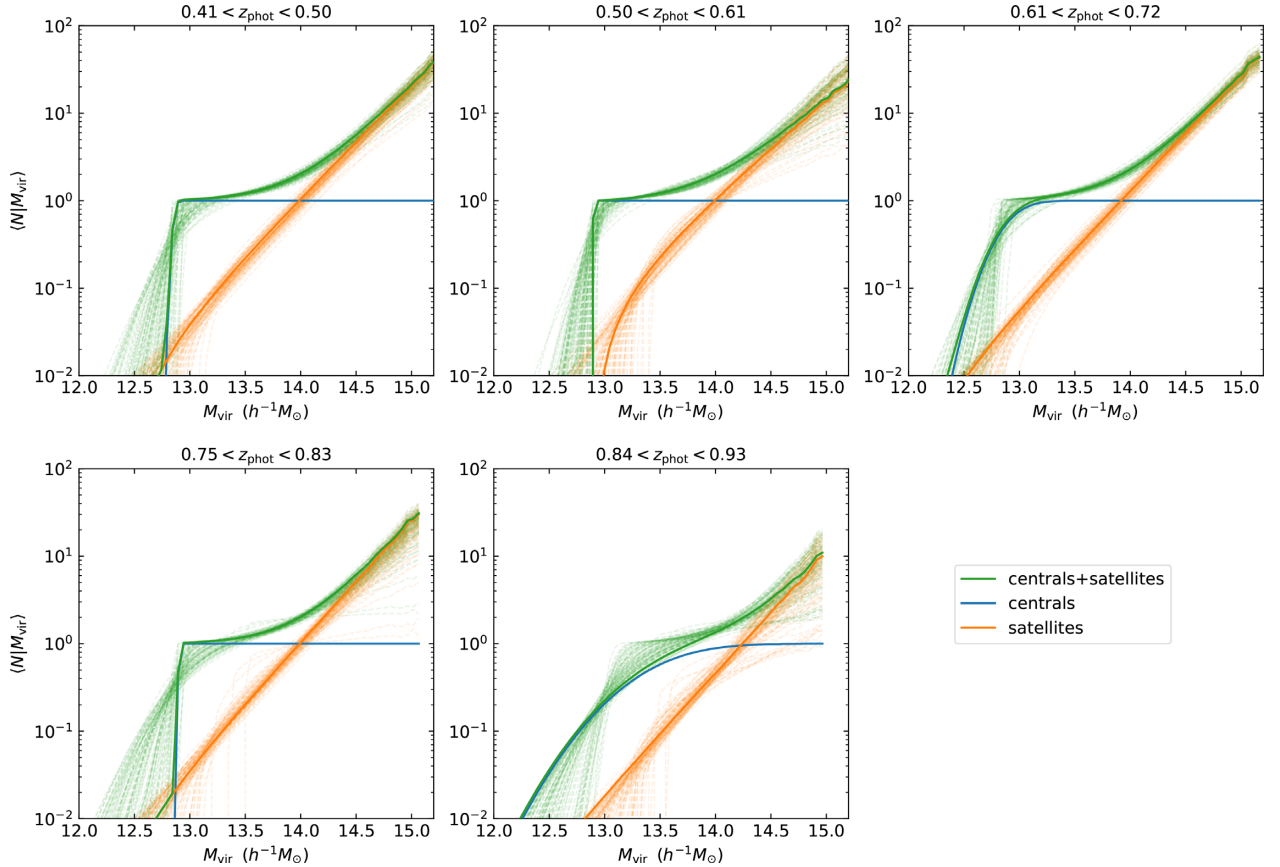


Figure 13. The halo occupations, i.e. the average number of galaxies that are hosted by a halo of a certain mass. The occupations of centrals, satellites, and the full occupation are plotted separately. The solid lines are from the best-fitting parameters. The dashed lines are from 100 sets of parameters randomly selected from the MCMC chain (only the satellite and full occupations are plotted for clarity).

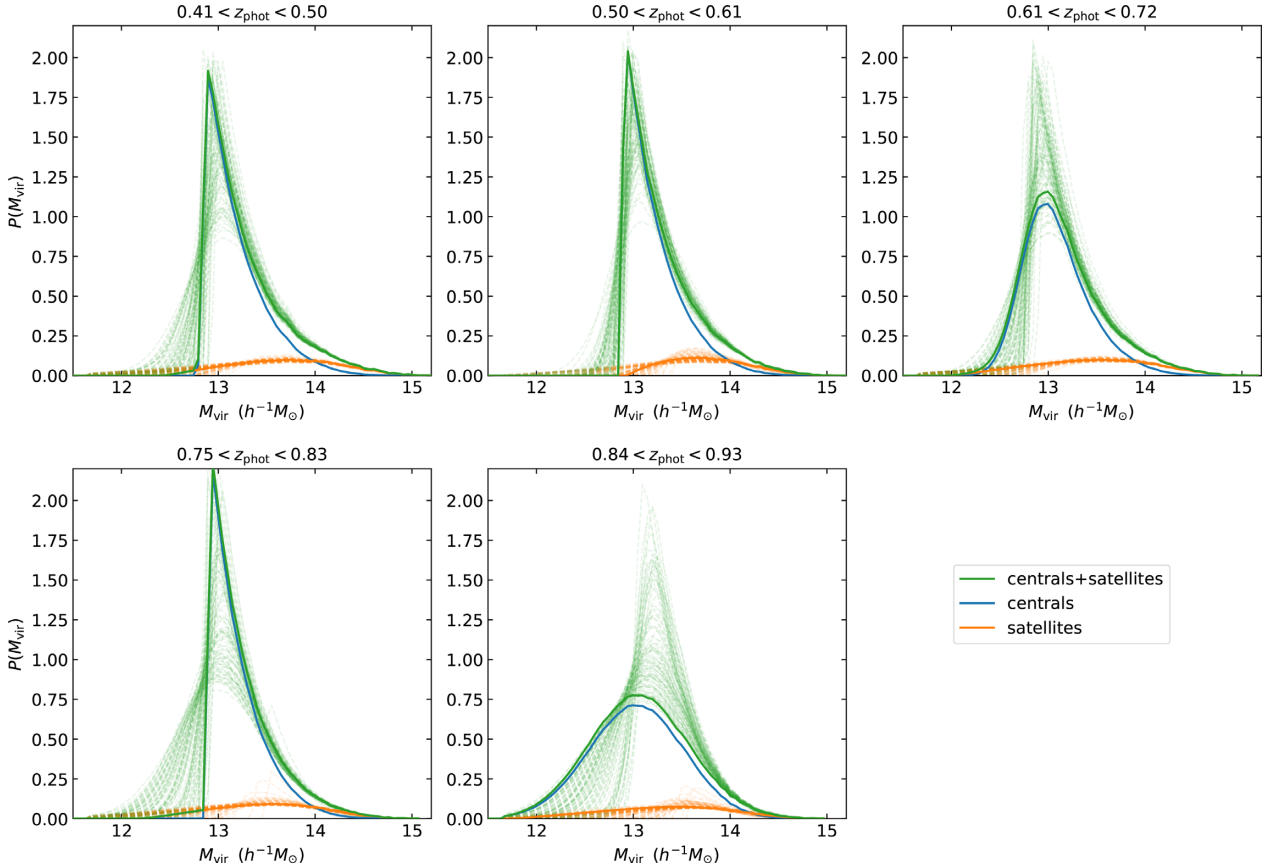


Figure 14. Similar to Fig. 13, but showing the probability distribution of host halo mass for a randomly selected galaxy in each redshift bin, with the solid lines showing the best-fitting results and dashed lines showing parameters from the MCMC chain. The green curve shows the host halo mass distribution for all galaxies (centrals + satellites) normalized as the probability per $\log M$; the blue and orange curves show the central and satellite components. The centrals in the highest-redshift sample have a much smoother low-mass cut-off than in the lower redshift samples.

N -body simulation as below

$$b = (\xi_{\text{gal}}/\xi_{\text{matter}})^{1/2}, \quad (10)$$

where ξ_{gal} and ξ_{matter} are the two-point correlation functions of galaxies and matter, respectively. (The galaxy bias is a function of scale, and at large scale it asymptotes to the large-scale bias value which we compute here; hereafter, we refer to the large-scale bias simply as the galaxy bias and denote it as b .) However, due to the lack of access to the MDPL2 dark matter particle catalogue, we instead use the analytical halo bias–halo mass relation from Tinker et al. (2010), implemented by the COLOSSUS software package (Diemer 2018). Specifically, we obtain the galaxy bias by averaging over the bias of all haloes weighted by the number of galaxies in each halo

$$b_{\text{gal}} = \frac{1}{N_{\text{gal}}} \sum_i b_{\text{halo}}(M_{\text{vir},i}) N_{\text{gal}}(i), \quad (11)$$

where N_{gal} is the total number of galaxies, $M_{\text{vir},i}$ is the virial mass of the i -th halo, and $N_{\text{gal}}(i)$ is the number of galaxies (central and satellites) in the i -th halo.

There is a strong degeneracy between S_z and galaxy bias, as shown in Fig. 15, which leads to large uncertainties in galaxy bias at high redshift. This degeneracy is expected, since a stronger clustering signal could be due to either a higher bias or smaller photo- z errors.

The evolution of galaxy bias with redshift, shown in Fig. 16, is consistent with the galaxy bias evolution of a sample with constant

clustering amplitude, and it can be written as $b(z) = 1.5/D(z)$, where $D(z)$ is the linear growth factor. The factor 1.5 is slightly smaller than the factor 1.7 assumed in the DESI Final Design Report (DESI Collaboration 2016b).

The BOSS CMASS sample (Dawson et al. 2013), which has a median redshift of $z \simeq 0.55$, was selected with similarly motivated luminosity threshold cuts to yield roughly half the comoving number density of our LRG sample. At $z = 0.55$, our galaxy bias estimates is consistent with $b \simeq 2.0$ from White et al. (2011) for the CMASS sample. The satellite fraction of our DESI-like LRGs is roughly 15 per cent, compared to 10 per cent for CMASS. The fact that there are much fewer satellite galaxies than central galaxies in LRG-like samples is a result of the selection cuts: only the most luminous galaxies are selected, and these luminous galaxies are much more likely to be at the centres of dark matter haloes.

The literature on HOD analysis for comparable LRGs at higher redshifts is relatively scarce. The eBOSS survey targeted LRGs in the redshift range of $0.6 < z < 0.9$ (Prakash et al. 2016), and Zhai et al. (2017) performed an HOD analysis on the combined BOSS + eBOSS sample in this redshift range. However, the eBOSS LRG sample is significantly different from our DESI-like LRGs in certain aspects: (1) the comoving number density of the DESI-like LRGs is more than 5 times that of the eBOSS LRGs, and (2) the eBOSS LRG selection does not contain luminosity-threshold cuts, resulting in a wider range of luminosity. Therefore, one should not expect the two samples to have the same HOD or derived parameters. The galaxy

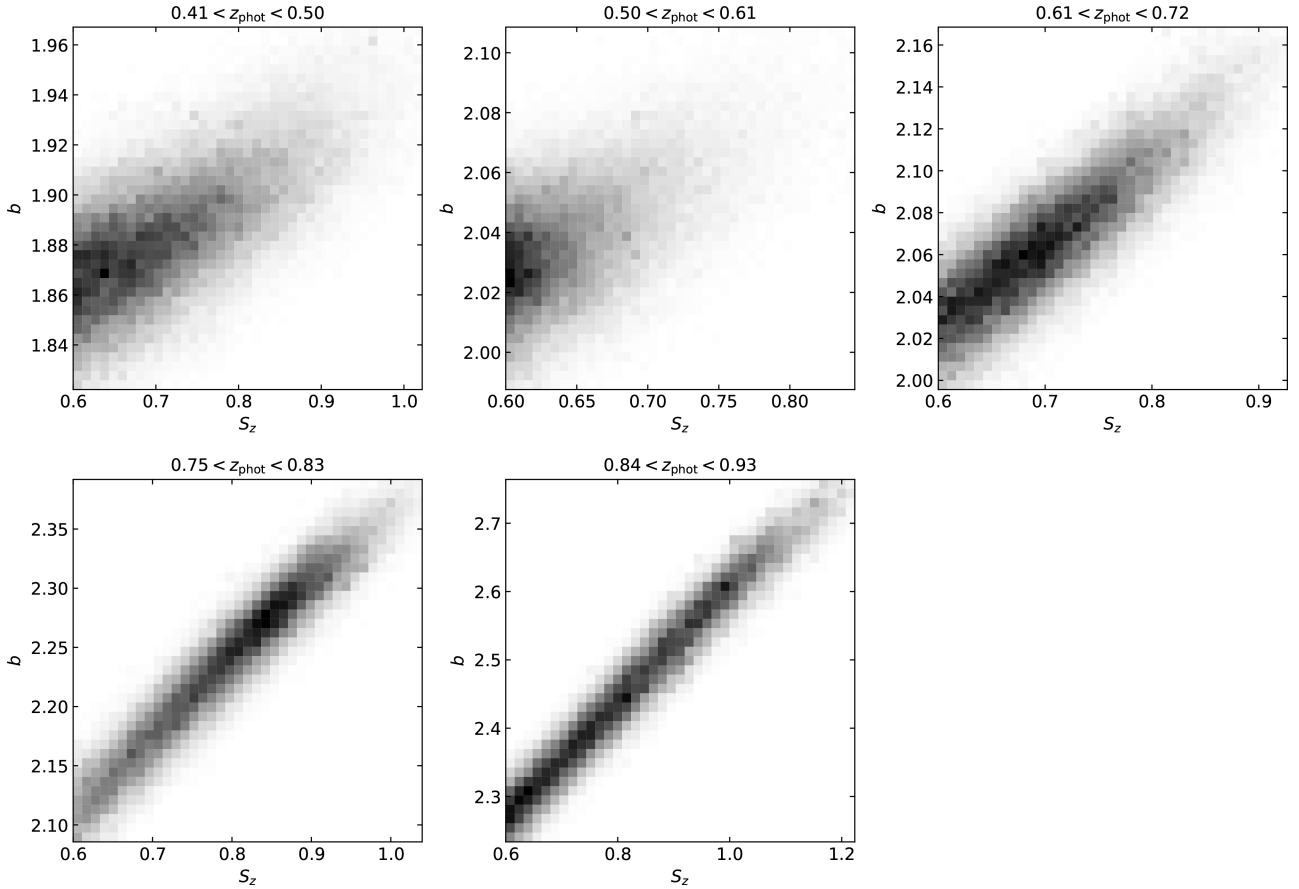


Figure 15. Two-dimensional posterior distribution of galaxy bias versus the photo- z error rescaling factor S_z (a nuisance parameter in our analysis) at different redshifts. The correlations are much stronger at higher redshifts due to the overall larger photo- z errors that dilute the clustering signal.

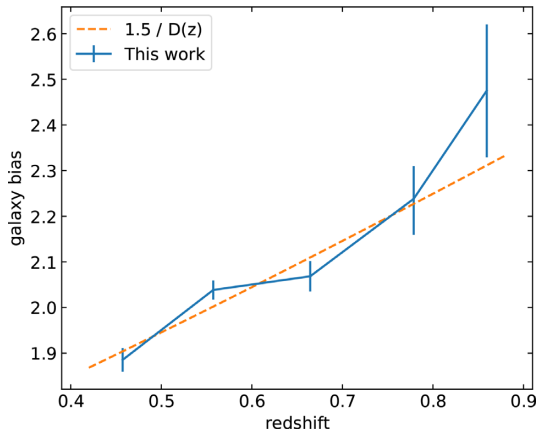


Figure 16. The evolution of galaxy bias for our LRG samples. The error bars show the 16 per cent and 84 per cent percentiles. The trend is consistent with the bias evolution that would be obtained if one assumed constant clustering amplitude, as shown in the dashed line.

bias of DESI-like LRGs at $z \sim 0.7$ (median redshift of eBOSS) is ~ 2.15 , compared to 2.3 for eBOSS LRGs. The DESI-like LRGs have a satellite fraction similar to the 13 per cent for eBOSS LRGs. The DESI-like LRGs have much smaller scatter in the halo mass threshold ($\sigma_{\log M}$) compared to eBOSS (which has $\sigma_{\log M} = 0.82$). Lower values of $\sigma_{\log M}$ for our LRG sample is expected since the sample selection includes a luminosity-threshold cut (whereas the

eBOSS LRG selection does not), but as we discussed earlier, our estimate might suffer from inflexibility in the HOD model or photo- z systematics.

The α parameter shows little variation with redshift and is slightly larger than unity; this value is roughly consistent with SDSS (Zehavi et al. 2011; Zentner et al. 2019) and BOSS (White et al. 2011) results. It is significantly larger than the value from Zhai et al. (2017) which estimated $\alpha \sim 0.4$, although this difference could be attributed to the aforementioned differences in the sample selection.

Since the HOD model is probabilistic, even if the HOD parameters are fixed, each HOD realization yields a different set of mock galaxies and thus slightly different clustering statistics. This effectively adds a noise to the likelihood function in the MCMC. So long as this noise has mean of zero, it can be shown that as the number of steps becomes large the distribution of points in the chain should still converge to the correct posterior. However, this ‘realization noise’ does cause the likelihoods associated with each step of the MCMC chain to have values which are biased high (or, equivalently, χ^2 to be biased low), since each ‘walker’ is less likely to move away from a point whose likelihood was evaluated to be higher than average, and more likely to move away from one which fluctuated low. We show the impact of this bias in Appendix D. When the realization noise is significant (as here), one cannot directly use the likelihood values from the chain to find the best-fitting point or to assess its χ^2 ; instead, it is necessary to average over repeated realizations of the same model parameters. This could be done at every step of the chain to reduce realization noise, but that is computationally expensive; instead, we adopt an alternative approach.

Specifically, we can exploit the fact that even though the likelihood value associated with each point in the chain is noisy and biased, the set of positions in parameter space that make up the chain do converge to match sampling from the posterior distribution. As a result, the density of points in the chain is highest where the posterior probability is greatest, even when the likelihood values assigned to those points may be inaccurate. We therefore select a small set of steps from the chains which lie in the highest density region of the parameter space; this set is highly likely to contain the points in the chain closest to the best-fitting parameters. We then compute the χ^2 values for each of these sets of parameters, averaging over a large number of realizations; from this we can find the point in the chains which truly has the highest likelihood. We describe this procedure in more detail in Appendix D.

Table 4 lists the best-fitting parameters and the corresponding averaged χ^2 values averaged over 1000 realizations. The table also lists the one-sided p -values corresponding to each χ^2 value; i.e. the probability of observing a χ^2 larger than the observed value purely by chance (if we find $p < 0.05$, the hypothesis that the best-fitting HOD model matches the data should be rejected). We compute this p -value using the number of degrees of freedom $N_{\text{dof}} = N_{\text{data}} - N_{\text{param}} = 13 - 6 = 7$, where N_{data} is the number of r_p bins plus one additional constraint from the comoving number density, and N_{param} is the number of free parameters. In every case, the HOD model returns a satisfactory fit.

7 DISCUSSION AND CONCLUSION

We have made a number of methodological improvements for galaxy clustering analysis with the HOD model using photo- z 's. We have developed a method that divides the irregular survey footprint into uniform subregions that allowed us to apply the jackknife resampling technique on this data set. Our methods of correlation measurements using the projected correlation function and the 'cross-correlation' L-S estimator recover many galaxy pairs straddling the boundaries of the redshift bins, and prevent the counting of pairs which are too far apart in photo- z to have significant clustering but are still placed in the same redshift bin. Both effects boost the S/N in clustering measurements compared to a purely angular clustering analysis. The methods also allow for straightforward and consistent modelling by assigning photo- z errors from the estimated error distribution to the mock galaxies.

With these aforementioned improvements, we have demonstrated that it is possible to obtain good constraints on HOD parameters using only photometric data. Specifically, we have measured the clustering and performed an HOD analysis for DESI LRG target galaxies. We have found that the LRGs are found in massive haloes (and especially so for high- z LRGs); this is expected since these are massive and red galaxies, which are only found in the densest environments (see e.g. Blanton & Moustakas 2009). We have also found that the host halo properties are very similar for all except the highest-redshift bin. The galaxy bias steadily increases with redshift, increasing from $b \simeq 1.9$ at $z \simeq 0.45$ to $b \simeq 2.3$ at $z \simeq 0.9$. This trend can be approximated by $b = 1.5/D(z)$, implying constant clustering amplitude over time. The fits prefer a relatively small scatter in the halo mass threshold, suggesting that the LRG selection is efficient in selecting galaxies in massive dark matter haloes. At high redshift, the host haloes are significantly more massive; this is due to the selection effect of the apparent magnitude limit on the galaxy sample.

The results of this paper can be used to create improved mock galaxy catalogues for DESI. The upcoming spectroscopic data from DESI will eliminate the uncertainties from photo- z 's and provide

tests of our results, although there will instead be systematics from fibre collisions to be dealt with. The spectroscopic redshifts will also enable the accurate measurements of the rest-frame colours and luminosity, and it would be interesting to study the colour and luminosity dependence of the galaxy clustering; results from such studies can provide important insights into the formation and evolution of these massive galaxies.

The overall methodology of HOD modelling with photo- z 's presented here can be easily implemented with existing analysis codes such as HALOTOOLS, and it can be adopted for future imaging surveys such as LSST for studying the galaxy-halo connection. There are several aspects where our methods can be further improved upon by adding more sophistication, which we discuss below.

First, the approximation of Gaussian photo- z errors is not always appropriate. In our case, the LRGs have prominent spectral features such as the 4000 Å; break and the 1.6 micron bump that result in unambiguous photo- z 's. Therefore, we are able to treat each PDF as a simple Gaussian distribution, and to assume that the photo- z errors are dominated by photometric uncertainties; these simplifications are supported by spectroscopic validation. However, in many other cases, the Gaussian approximation is not sufficient, and one needs to take the full photo- z PDFs as input in the fitting process; this can be important for galaxies that have skewed or multimodal PDFs or for data sets that have weak constraining power on redshifts.

Second, although in our case the uncertainty in the calibration of photo- z errors (quantified by S_z) is subdominant at lower redshifts, at higher redshift it causes significant uncertainties in HOD parameters and galaxy bias. Therefore, in the presence of relatively large photo- z errors, better priors on the calibration of photo- z uncertainties would significantly reduce the uncertainties in the model parameters. This can be achieved using a small spectroscopic subsample that is representative of the full photometric sample.

Third, in this work we have assumed that the our galaxy sample has the same intrinsic photo- z error properties, and the distribution of actual photo- z errors is solely due to the variation in the S/N of the photometry; therefore, we can randomly draw from the estimated photo- z errors and assign the resulting values to each mock galaxy. This assumption might not hold for a sample of galaxies that are more diverse than the LRGs; for example, in a pure luminosity threshold sample, the blue and red galaxies will have very different intrinsic photo- z errors/PDFs. Nevertheless, the method can account for such differences by treating differently haloes with different properties when assigning photo- z errors to the corresponding galaxies.

Finally, since the correlation function is measured with relative distances, it is insensitive to an overall offset in photo- z 's so long as the offset is the same for all galaxies in the sample. Our model does not account for higher order offsets, and light-cone mocks would be required to simulate such effects.

With these improved methods and enlarged samples from future surveys, fully photometric HOD modelling will be a powerful tool for studying the galaxy-halo connection with future imaging surveys such as LSST.

ACKNOWLEDGEMENTS

The authors would like to thank Hee-Jong Seo, Jeremy Tinker, and Gustavo Niz for their feedback on the draft and useful discussions. RZ and JAN were supported by the U.S. Department of Energy Office of Science, Office of High Energy Physics via grant DE-SC0007914. RZ also is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC02-05CH1123. Support for YYM was provided by the

Pittsburgh Particle Physics, Astrophysics and Cosmology Center through the Samuel P. Langley PITT PACC Postdoctoral Fellowship, and by NASA through the NASA Hubble Fellowship grant no. *HST*-HF2-51441.001 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555. JM gratefully acknowledges support from NSF grant AST-1616414 and DOE grant DE-SC0020086. ADM was supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award Number DE-SC0019022. ARZ was funded by the US National Science Foundation (NSF) through grants AST 1516266 and AST 1517563. DYT thanks Prof. Steve Ahlen for his mentorship and support and acknowledges the generous support by U.S. Department of Energy Office of Science, grant No. DE-SC0015628.

This research is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC02-05CH1123, and by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract; additional support for DESI is provided by the U.S. National Science Foundation, Division of Astronomical Sciences under Contract No. AST-0950945 to the National Optical Astronomy Observatory; the Science and Technologies Facilities Council of the United Kingdom; the Gordon and Betty Moore Foundation; the Heising-Simons Foundation; the French Alternative Energies and Atomic Energy Commission (CEA); the National Council of Science and Technology of Mexico, and by the DESI Member Institutions. The authors are honoured to be permitted to conduct astronomical research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation.

The Legacy Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS; NOAO Proposal ID # 2014B-0404; PIs: David Schlegel and Arjun Dey), the Beijing-Arizona Sky Survey (BASS; NOAO Proposal ID # 2015A-0801; PIs: Zhou Xu and Xiaohui Fan), and the Mayall z-band Legacy Survey (MzLS; NOAO Proposal ID # 2016A-0453; PI: Arjun Dey). DECaLS, BASS, and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory (NOAO); the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOAO. The Legacy Surveys project is honoured to be permitted to conduct astronomical research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation.

NOAO is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

This project used data obtained with the Dark Energy Camera (DECam), which was constructed by the Dark Energy Survey (DES) collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo, Financiadora de Estudos e Projetos, Fundacao Carlos

Chagas Filho de Amparo a Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Cientifico e Tecnológico and the Ministerio da Ciencia, Tecnologia e Inovacao, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zurich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

BASS is a key project of the Telescope Access Program (TAP), which has been funded by the National Astronomical Observatories of China, the Chinese Academy of Sciences (the Strategic Priority Research Program 'The Emergence of Cosmological Structures' Grant # XDB09000000), and the Special Fund for Astronomy from the Ministry of Finance. The BASS is also supported by the External Cooperation Program of Chinese Academy of Sciences (Grant # 114A11KYSB20160057), and National Natural Science Foundation of China (Grant # 11433005). The Legacy Survey team makes use of data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), which is a project of the Jet Propulsion Laboratory/California Institute of Technology. NEOWISE is funded by the National Aeronautics and Space Administration.

The Legacy Surveys imaging of the DESI footprint is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC02-05CH1123, by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract; and by the U.S. National Science Foundation, Division of Astronomical Sciences under Contract No. AST-0950945 to NOAO.

Funding for the DEEP2 Galaxy Redshift Survey has been provided by NSF grants AST-95-09298, AST-0071048, AST-0507428, and AST-0507483 as well as NASA LTSA grant NNG04GC89G. Funding for the DEEP3 Galaxy Redshift Survey has been provided by NSF grants AST-0808133, AST-0807630, and AST-0806732. Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on data taken from the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the

GAMA regions is being obtained by a number of independent survey programmes including GALEX MIS, VST KiDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT and ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO, and the participating institutions. The GAMA website is <http://www.gama-survey.org/>. This paper uses data from the VIMOS Public Extragalactic Redshift Survey (VIPERS). VIPERS has been performed using the ESO Very Large Telescope, under the ‘Large Programme’ 182.A-0886. The participating institutions and funding agencies are listed at <http://vipers.inaf.it>. This research uses data from the VIMOS VLT Deep Survey, obtained from the VVDS database operated by Cesam, Laboratoire d’Astrophysique de Marseille, France.

DATA AVAILABILITY

The LRG catalogue and the photometric redshifts are derived from the publicly available Legacy Surveys imaging data.⁵ The LRG catalogue, masks, and the specific version of the photo- z ’s used in the clustering analysis will be shared on reasonable request to the corresponding author.

REFERENCES

- Abolfathi B. et al., 2018, *ApJS*, 235, 42
 Aihara H. et al., 2018a, *PASJ*, 70, S8
 Aihara H. et al., 2018b, *PASJ*, 70, S4
 Alam S. et al., 2017, *MNRAS*, 470, 2617
 Baldry I. K. et al., 2018, *MNRAS*, 474, 3875
 Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, *ApJ*, 762, 109
 Berlind A. A., Weinberg D. H., 2002, *ApJ*, 575, 587
 Blake C. et al., 2016, *MNRAS*, 462, 4240
 Blanton M. R., Moustakas J., 2009, *ARA&A*, 47, 159
 Breiman L., 2001, *Mach. Learn.*, 45, 5
 Childress M. J. et al., 2017, *MNRAS*, 472, 273
 Cooper M. C. et al., 2012, *MNRAS*, 419, 3018
 Cutri R. M. et al., 2013, Technical Report, Explanatory Supplement to the AllWISE Data Release Products
 Dahlen T. et al., 2013, *ApJ*, 775, 93
 Dawson K. S. et al., 2013, *AJ*, 145, 10
 Dawson K. S. et al., 2016, *AJ*, 151, 44
 DESI Collaboration, 2016a, preprint ([arXiv:1611.00037](https://arxiv.org/abs/1611.00037))
 DESI Collaboration, 2016b, preprint ([arXiv:1611.00036](https://arxiv.org/abs/1611.00036))
 Dey A. et al., 2019, *AJ*, 157, 168
 Diemer B., 2018, *ApJS*, 239, 35
 Eisenstein D. J. et al., 2005, *ApJ*, 633, 560
 Foreman-Mackey D., 2016, *J. Open Source Softw.*, 1, 24
 Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
 Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, *ApJ*, 622, 759
 Gunn J. E. et al., 1998, *AJ*, 116, 3040
 Hartigan J. A., Wong M. A., 1979, *Appl. Stat.*, 28, 100
 Hearin A. et al., 2017, *AJ*, 154, 190
 Hodges J. L., Lehmann E. L., 1963, *Ann. Math. Stat.*, 34, 598
 Høg E. et al., 2000, *A&A*, 355, L27
 Jarvis M., Bernstein G., Jain B., 2004, *MNRAS*, 352, 338
 Kind M. C., Brunner R. J., 2013, *MNRAS*, 432, 1483
 Klypin A., Yepes G., Gottlober S., Prada F., Hess S., 2016, *MNRAS*, 457, 4340
 Kochanek C. S. et al., 2012, *ApJS*, 200, 8
 Laigle C. et al., 2016, *ApJS*, 224, 24
 Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64

- Lang D., 2014, *AJ*, 147, 108
 Lang D., Hogg D. W., Mykytyn D., 2016, Astrophysics Source Code Library, record ascl:1604.008
 Le Fevre O. et al., 2013, *A&A*, 559, A14
 Malz A. I., Hogg D. W., 2020, preprint ([arXiv:2007.12178](https://arxiv.org/abs/2007.12178))
 Meisner A. M., Lang D., Schlafly E. F., Schlegel D. J., 2019, *PASP*, 131, 124504
 Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
 Newman J. A. et al., 2013, *ApJS*, 208, 5
 Parkinson D. et al., 2012, *Phys. Rev. D*, 86
 Pedregosa F. et al., 2011, *J. Mach. Learn. Res.*, 12, 2825
 Planck Collaboration XVI, 2014, *A&A*, 571, A16
 Prakash A., Licquia T. C., Newman J. A., Rao S. M., 2015, *ApJ*, 803, 105
 Prakash A. et al., 2016, *ApJS*, 224, 34
 Scodreggio M. et al., 2018, *A&A*, 609, A84
 Sinha M., Garrison L., 2017, Astrophysics Source Code Library, record ascl:1703.003
 Skrutskie M. F. et al., 2006, *AJ*, 131, 1163
 Soo J. Y. H. et al., 2018, *MNRAS*, 475, 3613
 Strauss M. A. et al., 2002, *AJ*, 124, 1810
 Tanaka M. et al., 2018, *PASJ*, 70, S9
 Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlober S., 2010, *ApJ*, 724, 878
 The Dark Energy Survey Collaboration, 2005, preprint ([astro-ph/0510346](https://arxiv.org/abs/astro-ph/0510346))
 Wechsler R. H., Tinker J. L., 2018, *ARA&A*, 56, 435
 White M. et al., 2011, *ApJ*, 728, 126
 Wright E. L. et al., 2010, *AJ*, 140, 1868
 York D. G., 2000, *AJ*, 120, 1579
 Zehavi I. et al., 2011, *ApJ*, 736, 59
 Zentner A. R., Hearin A., van den Bosch F. C., Lange J. U., Villarreal A., 2019, *MNRAS*, 485, 1196
 Zhai Z. et al., 2017, *ApJ*, 848, 76
 Zheng Z., Coil A. L., Zehavi I., 2007, *ApJ*, 667, 760
 Zheng Z., Zehavi I., Eisenstein D. J., Weinberg D. H., Jing Y. P., 2009, *ApJ*, 707, 554
 Zhou R. et al., 2019, *MNRAS*, 488, 4565

APPENDIX A: WISE BRIGHT STAR MASKS

This appendix provides some technical details of the *WISE* bright star masks that we described in Section 2.2. The *WISE* bright star masks are geometric masks around stars in the AllWISE catalogue with $W1_{AB} < 16.0$. The masks consist of two components: circular masks for the ‘core’ of a bright star, and rectangular masks for the diffraction spikes. The size of the masks vary with the $W1$ magnitude of the

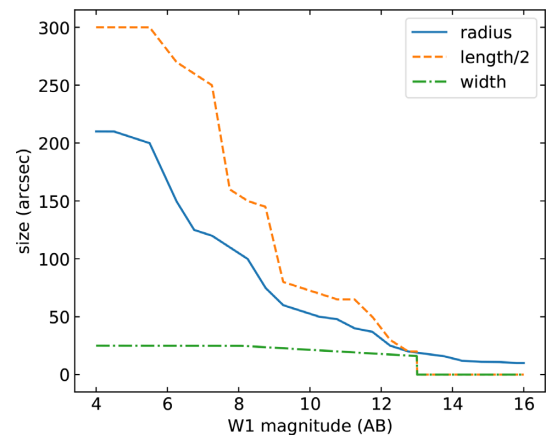


Figure A1. The circular mask radius and the length and width of the rectangular mask as a function of $W1$ magnitude of the bright star. No masking of diffraction spikes is performed for stars fainter than $W1_{AB} = 13.0$.

⁵<https://www.legacysurvey.org/>

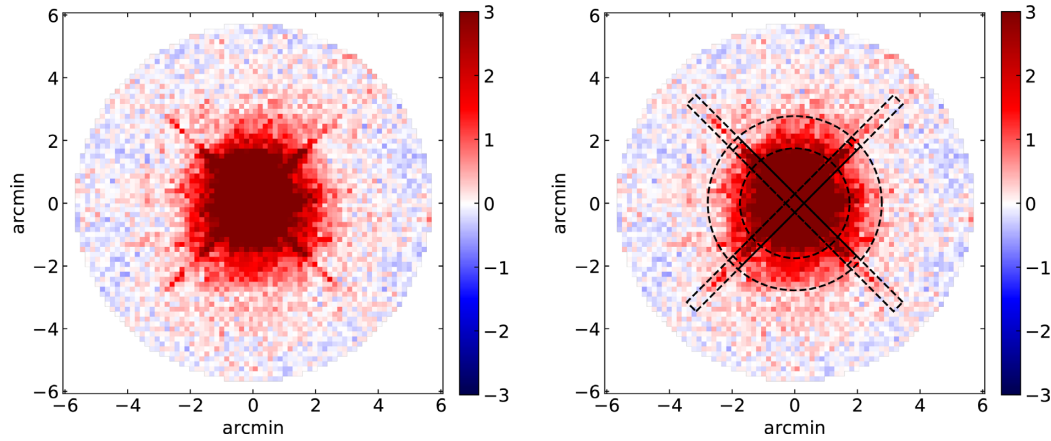


Figure A2. Left-hand panel: Cross-correlation between LRGs and AllWISE bright stars with $6 < W1_{AB} < 8$ in ecliptic coordinates. The colours represent the fractional over/underdensity of LRGs compared to the overall average density (as a result the values cannot drop below -1). Right-hand panel: The boundaries of the WISE masks at $W1_{AB} = 6$ and $W1_{AB} = 8$ are overlaid.

bright star, and the size–magnitude relation is shown in Fig. A1. To obtain the size–magnitude relation, we cross-correlate the positions on the sky between LRGs and bright stars in magnitude bins of $\Delta W1_{AB} = 0.5$, and locate (by hand) the distance where the LRG density starts to noticeably deviate from the densities further away from the stars. This way the majority of the contaminated objects in the LRG sample are masked. Fig. A2 shows the cross-correlation between the LRG sample and AllWISE stars with $6 < W1_{AB} < 8$ with the masks overlaid.

APPENDIX B: PHOTOMETRIC REDSHIFTS FOR LEGACY SURVEYS DATA RELEASE 8

Accompanying this paper, we have released an updated version of the photometric redshifts used to conduct the large-scale structure analyses we have presented. This constitutes the Photometric Redshifts for the Legacy Surveys (PRLS) catalogue; it is built using the most recent DR8 data set of the Legacy Surveys.⁶ Data Release 8 includes data from both the DECaLS survey in the southern portion of the planned DESI footprint ($\text{Dec.} \leq 34^\circ$) and the BASS and MzLS surveys in the northern sky ($\text{Dec.} \geq 32^\circ$). The two regions have slightly different effective response curves due to variations in filter transmission, detectors (for the g and r bands), etc., and thus photo- z algorithms for each must be trained separately. Table B1 lists the redshift surveys used as training samples in each region. In addition to the redshift catalogues used for training the DR7 photo- z 's (described in Section 3.2), in the BASS/MzLS region we can also utilize redshifts from the DEEP3 Galaxy Redshift Survey (Cooper et al. 2012; Zhou et al. 2019). The method for computing the DR8 photo- z 's is the same as described in Section 3; as a result the photo- z performance in DR8 South is very similar to the DR7 photo- z 's used in this paper.

The columns of the PRLS photo- z catalogue are described in Table B2.

To assess the photo- z accuracy for the spectroscopic training sample, we employ five-fold cross-validation. To do this, we randomly divide the data set into five equal chunks; we can then combine four chunks for training and evaluate the performance with the remaining chunk. We repeat this procedure until all five chunks have been used

Table B1. Number of objects from each redshift survey that are cross-matched to Legacy Surveys DR8 (before downsampling). For data sets that have been downsampled for photo- z training, we also list percentages of the remaining objects after downsampling.

Survey	DR8 North	DR8 South
2dFLenS	–	28 416
AGES	16 197	4057
BOSS	377 769 (39.6 per cent)	951 829 (33.3 per cent)
COSMOS2015	–	62 510
DEEP2 Field 2, 3, 4	14 401	20 602
DEEP2 Field 1 +	13 355	–
DEEP3	–	123 293 (52.9 per cent)
GAMA	–	5174
OzDES	244	5174
SDSS	281 636 (46.0 per cent)	574 767 (43.9 per cent)
VIPERS	–	47 328
VVDS	–	6321
WiggleZ	1627	144 963 (32.6 per cent)
eBOSS	31 330	38 459

for testing. In this way, the entire truth data set can be utilized for testing without biasing the assessment of performance.

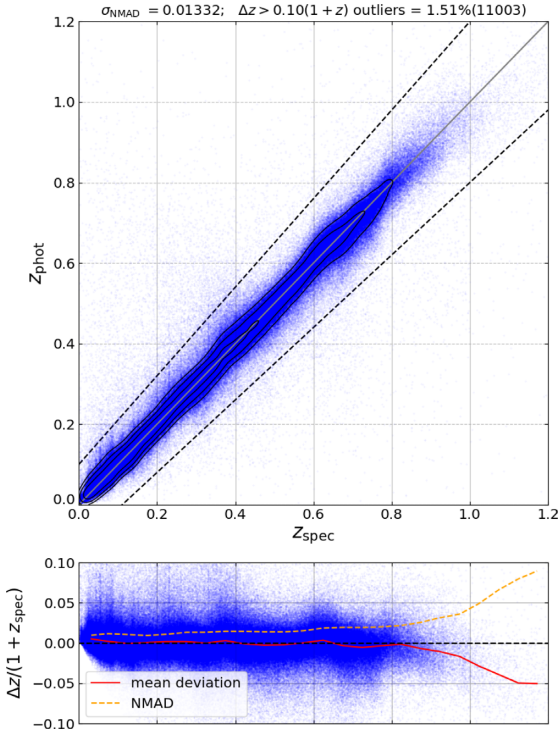
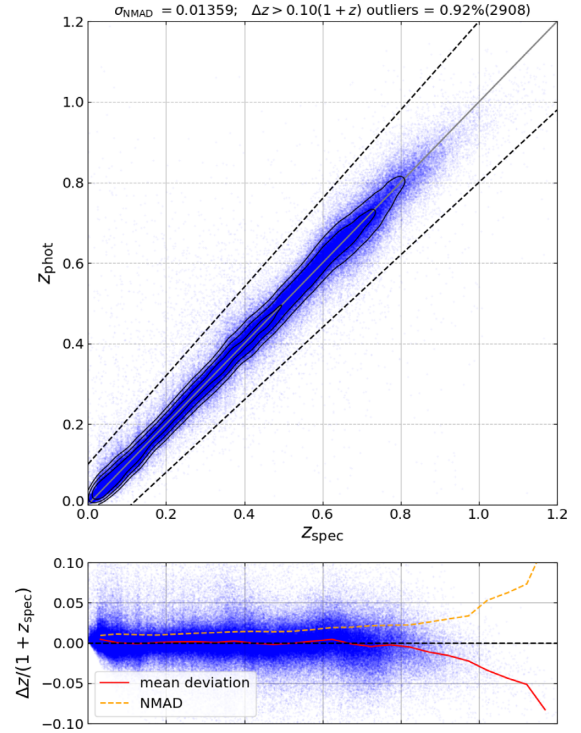
Fig. B1 shows the relationship between photo- z and spec- z for $z_{\text{mag}} < 21.0$ objects in the truth catalogue in DR8 South (the DECaLS region). For the (unweighted) objects, the photo- z scatter is ~ 0.013 and the outlier rate is 1.5 per cent, although it is worth noting that this sample is dominated by bright galaxies from surveys like SDSS and BOSS and therefore the numbers do not represent the photo- z accuracy of a magnitude limited sample with $z_{\text{mag}} < 21.0$. Fig. B2 presents the equivalent plot for DR8 North (the BASS/MzLS region).

To show how the photo- z 's start to systematically break down beyond $z_{\text{mag}} \simeq 21$, we plot photo- z versus spec- z for $z_{\text{mag}} > 21.0$ objects (in DR8 South) in Fig. B3. The failure of the photo- z 's is due to the limitations in our imaging data: $W1$ and $W2$ are too shallow to be useful for the fainter galaxies, so for these galaxies we are effectively limited to only the three optical bands, grz , which are not sufficient to constrain the photo- z 's. The shallow $W1/W2$ imaging is particularly problematic for galaxies with $z \lesssim 0.5$, as their 1.6 micron bump is still far from the $W1$ band, causing their photo- z 's to fail catastrophically.

⁶<https://www.legacysurvey.org/dr8/>

Table B2. Descriptions of the columns in the PRLS photometric redshift catalogue.

Name	Description
$z_{\text{phot_mean}}$	Photo- z derived from the mean of the photo- z PDF
$z_{\text{phot_median}}$	Photo- z derived from the median of the photo- z PDF
$z_{\text{phot_std}}$	Standard deviation of the photo- z 's derived from the photo- z PDF
$z_{\text{phot_l68}}$	Lower bound of the 68 per cent confidence region, derived from the photo- z PDF
$z_{\text{phot_u68}}$	Upper bound of the 68 per cent confidence region, derived from the photo- z PDF
$z_{\text{phot_l95}}$	Lower bound of the 95 per cent confidence region, derived from the photo- z PDF
$z_{\text{phot_u95}}$	Upper bound of the 68 per cent confidence region, derived from the photo- z PDF
z_{spec}	Spectroscopic redshift, if available
Survey	Source of the spectroscopic redshift
Training	Whether or not the spectroscopic redshift is used in photometric redshift training

**Figure B1.** Photo- z versus spec- z plot similar to Fig. 4, but for $z_{\text{mag}} < 21.0$ objects with spec- z training objects in DR8 South (the region covered by the DECaLS survey) without any weighting (cross-validation is used to avoid biasing errors low). The photometric redshifts are generally well behaved in this regime.**Figure B2.** As Fig. B1, but for DR8 North (the BASS/MzLS survey region).

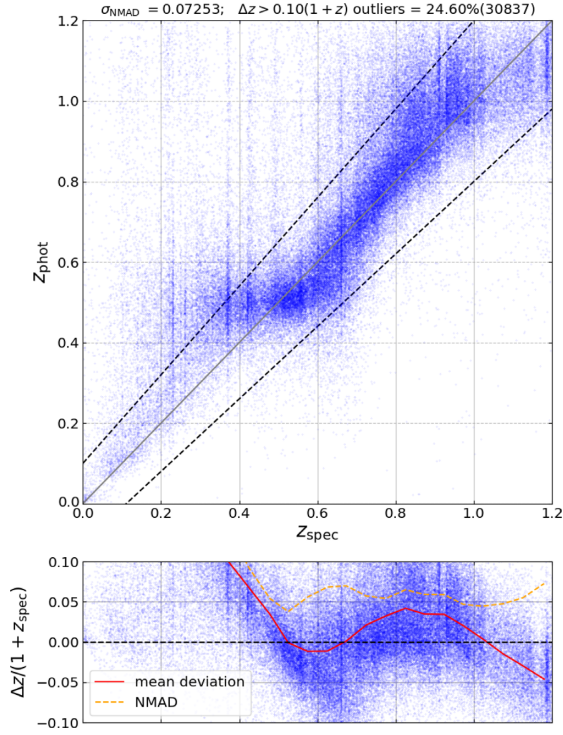


Figure B3. Same as Fig. B1 but for $z_{\text{mag}} > 21.0$ objects in the truth catalogue. In this regime, the photo- z 's are poorly constrained, especially for objects with $z \lesssim 0.5$.

APPENDIX C: HOD PARAMETERS FROM SKEW NORMAL PHOTO-Z PDF

The distributions of $(z_{\text{phot}} - z_{\text{spec}})/\sigma_z$ shown in Fig. 5 suggests that the distribution of photo- z errors may be slightly skewed (e.g. for the redshift bin covering $0.61 < z < 0.72$). To see if such a skewness could cause systematic offsets in our HOD modelling, we have rerun the MCMC for the redshift bin of $0.61 < z < 0.72$ treating all individual objects' photo- z PDFs as skew normal distributions with the same standard deviation that was used in the Gaussian

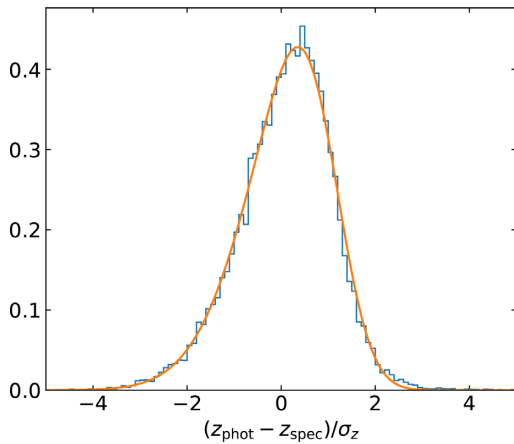


Figure C1. Illustration of our fit of a skew normal distribution to the redshift error distribution. As in Fig. 5, the histogram shows the distribution of redshift errors for objects with spectroscopic measurements in units of the estimated uncertainty for a given object for the $0.61 < z < 0.72$ redshift bin. The curve corresponds to the PDF of the best-fitting skew normal distribution, with $\alpha = -2.1$.

Table C1. The mean values and the 16th and 84th percentiles of the HOD parameters from the MCMC chain, for the Gaussian photo- z PDF and the skew normal PDF. In each case differences are small compared to the statistical uncertainties.

	$\log(M_{\text{min}})$	$\sigma \log M$	α	$\log(M_0)$	$\log(M_1)$
Gaussian	$12.86^{+0.05}_{-0.05}$	$0.20^{+0.13}_{-0.13}$	$1.31^{+0.06}_{-0.06}$	$11.94^{+0.58}_{-0.62}$	$13.92^{+0.04}_{-0.04}$
Skew normal	$12.86^{+0.05}_{-0.05}$	$0.20^{+0.13}_{-0.14}$	$1.31^{+0.06}_{-0.07}$	$11.98^{+0.56}_{-0.61}$	$13.92^{+0.04}_{-0.04}$

representation. We obtain the α parameter of the distribution (which determines its skewness) by fitting a skew normal distribution to the $(z_{\text{phot}} - z_{\text{spec}})/\sigma_z$ distribution, as shown in Fig. C1. We find the best-fitting α value to be -2.1 , which we keep fixed in the MCMC, using the same α value to skew the error distributions for each mock object. As in our standard analysis, the photo- z errors are rescaled by a constant factor of S_z which is treated as a nuisance parameter. The resulting HOD parameters are listed in Table C1, and for comparison we also list the results from our standard analysis (which treats the photo- z PDFs as Gaussian). We find that there are negligible changes in the inferred HOD parameters when switching to a skew normal PDF with the skewness estimated from the available spec- z data and standard deviation set by our error estimates.

APPENDIX D: IDENTIFYING MAXIMUM-LIKELIHOOD VALUES IN NOISY MCMC

Here, we provide more details on the treatment of noisy MCMC that we discussed in the last part of Section 6.

First, to demonstrate that the likelihoods (or, equivalently, χ^2 values, since we use flat priors and Gaussian likelihoods) at each step from the MCMC chain are indeed biased, we randomly select 100 points from the chain (where each point corresponds to a set of model parameters), and for each of these points we rerun 100 HOD realizations to obtain the average χ^2 value. Fig. D1 shows the difference between the χ^2 value taken from the MCMC chain and the mean χ^2 value averaging over 100 realizations for the same position in parameter space, drawn from the fits for the redshift bin of $0.61 < z < 0.72$ (the same redshift bin is used for all other figures in

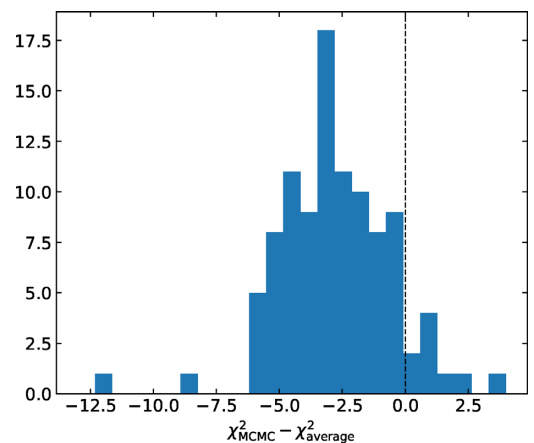


Figure D1. Histogram of the difference between the χ^2 value from the MCMC chain and the mean χ^2 value averaging over 100 realizations for the same set of HOD parameters. The vertical line highlights the location of zero difference. The χ^2 difference is clearly non-zero on average, indicating that the χ^2 values from the MCMC chain are biased low.

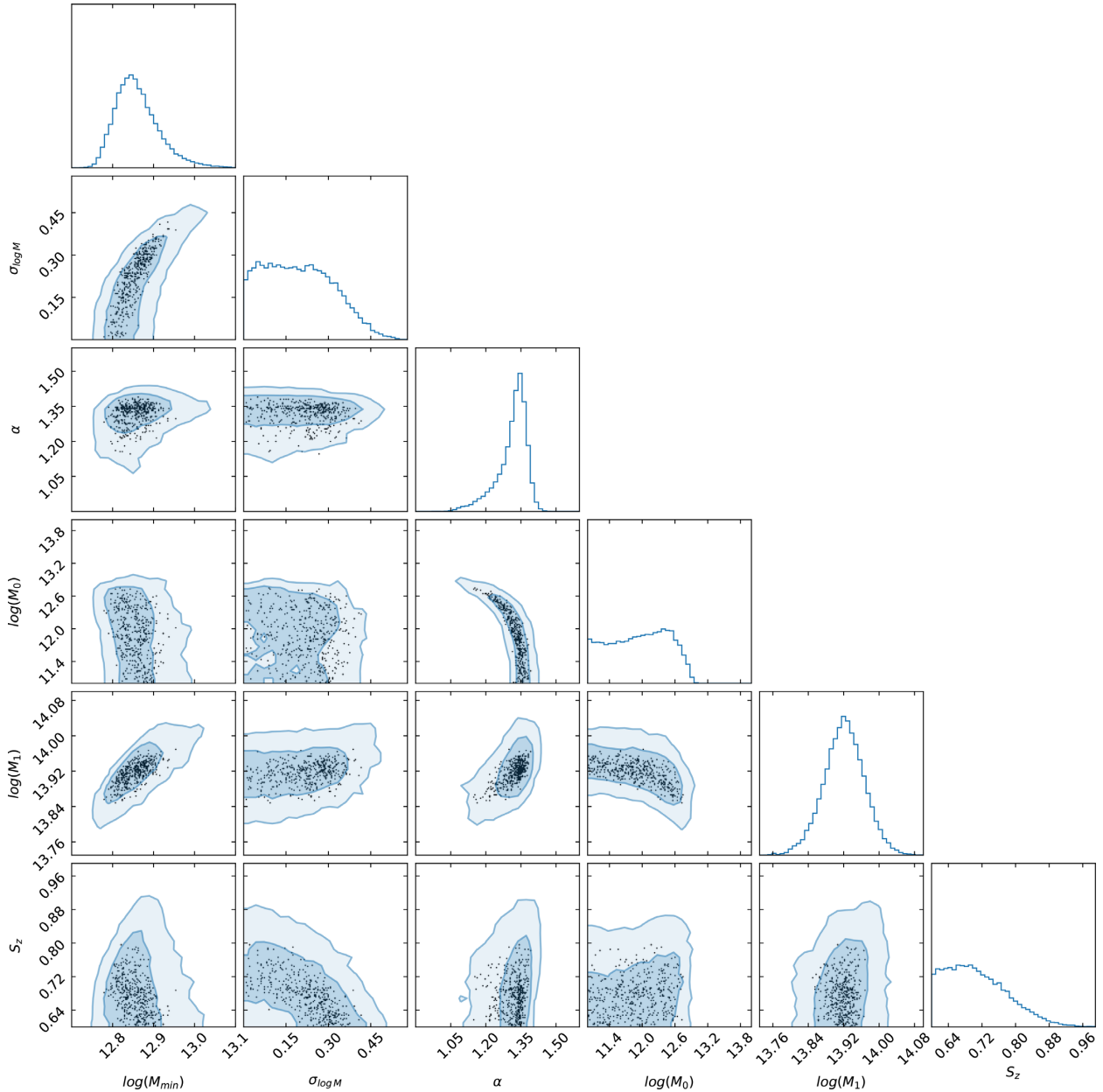


Figure D2. Points with the highest chain likelihood values selected from 500 segments in the chain for the redshift bin of $0.61 < z < 0.72$. The contours of the full MCMC chain are also shown. The highest-likelihood points span a wide range in parameter space due to the stochastic nature of the model likelihood evaluated at a given point in the chain.

this section). Clearly, the χ^2 values from the chain are biased low compared to the averaged values. This bias is caused by noise in the MCMC likelihood function, as discussed in Section 6.

To demonstrate the difficulty of using the likelihood from the MCMC chain for finding the best-fitting parameters in the presence of realization noise, here we show that the points selected to have the highest likelihood values have a very wide distribution in the parameter space. In MCMC, the chain positions are correlated, and sometimes a walker can get ‘stuck’ at the same position for many steps; to reduce such effects, we divide the chain into 500 segments, and select the highest likelihood point in each segment, and plot their positions in parameter space in Fig. D2. Even though these points have higher likelihoods in the chain than 99.95 per cent of the sample, they span a large range in parameter space.

As selecting the best-fitting parameters using the chain likelihoods does not yield good results, we adopt an alternative approach. The distribution of points in the chain should converge to follow the posterior. As a result, the density of points in the chain should be greatest where the posterior is highest. We can therefore select points in the regions of highest density and they are likely to be near the likelihood peak. Below we describe the method in detail.

First, we downsample the MCMC chain by selecting every 50th point from the chain to reduce the effects of correlations, and then for each point compute the distance to its 500th nearest neighbour (after downsampling). These distances should anticorrelate strongly with the local density. For the nearest neighbour search and distance calculation, the parameter space is normalized by the 16–84 per cent percentile range along each dimension. We also perform ‘reflection’

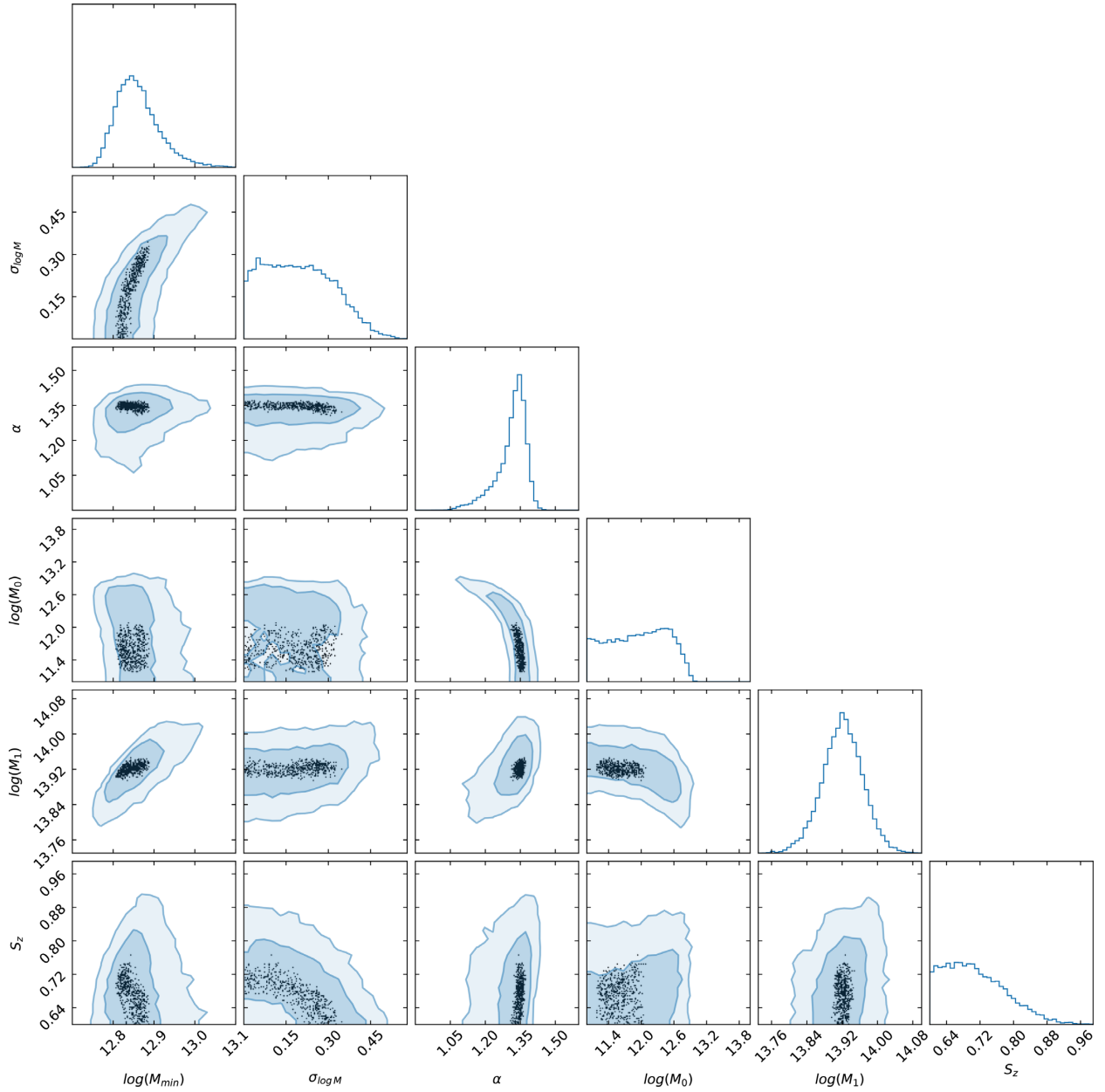


Figure D3. The distribution of 500 points that are in the highest-density region in parameter space. The contours of the full MCMC chain are also shown. The distribution of these points is much more compact than the distribution of points selected directly from MCMC likelihoods which is shown in the previous plot.

on $\sigma_{\log M}$ and S_z dimensions at their lower boundaries in their prior to remove the boundary effect: each point in the chain is duplicated with the same parameters except for $\sigma_{\log M}$ which adopts the value of $0 - \sigma_{\log M}$ where 0 is the lower boundary of $\sigma_{\log M}$; subsequently the same procedure is also performed for S_z , so in the end we have 4 times the original number of points.

We select the 500 points with the smallest neighbour distances. The distribution of these points is shown in Fig. D3; note that these points have a much more compact distribution than those in Fig. D2. For each of these points, we generate 100 HOD realizations and determine the χ^2 for each. We then compute the averaged χ^2 using

the Hodges–Lehmann estimator (Hodges & Lehmann 1963). Since there is still some scatter in the averaged χ^2 , we again select the 10 points with the smallest averaged χ^2 and generate 1000 HOD realizations to get more accurate χ^2 values. The point with the smallest averaged χ^2 is selected as the set of best-fitting parameters. In some redshift bins, the distributions of the 10 points are more compact than the 500 points; in other bins, the distributions are rather similar.

This paper has been typeset from a \LaTeX file prepared by the author.